Point dose measurements in VMAT: an investigation of detector choice and plan complexity.

Daniel Phillip Goodwin

Department of Physics and Astronomy

University of Canterbury

A thesis submitted in partial fulfilment of the requirements for the Degree of

Master of Science in Medical Physics

Auckland City Hospital March 2017

Acknowledgements

Firstly, I would like to express my gratitude to my supervisors James Talbot, Lydia Handsfield, and Dr. Steven Marsh. James, your guidance and continued support (even from the other side of the planet) have been invaluable. Lydia, I thank you for your willingness oversee the final stages of this project, your enthusiasm and excellent eye for detail have driven this write up.

To my good friends and colleagues Neil Campbell and Suzanne Lydiard, for your guidance, constant criticisms (Neil), and motivational support (Suzy), I thank you. I would like to acknowledge all the other members of the physics team at Auckland City Hospital for their encouragement and insightful comments.

I must take this opportunity to apologise to my fellow registrars and officemates for my aggressively loud sighing, pacing, and incoherent mumbling during the last few months (years).

Finally to my friends and family, thank you for keeping me relatively sane.

Abstract

Purpose: Volumetric Modulated Arc Therapy (VMAT) is an intensity modulated radiation therapy technique which can achieve highly conformal dose delivery through dynamic variation of dose rate, gantry speed, and multileaf collimator positions. Due to the complexity of treatment delivery, patient specific quality assurance (QA) is required to ensure agreement between calculated and delivered dose. Point dose measurements are a well established patient specific QA technique for VMAT. The aim of this study was to investigate the relationship between plan complexity and the agreement between measured and calculated point doses. The suitability of five different detectors for VMAT point dose measurements was also evaluated.

Methods: 45 previously treated prostate VMAT plans were selected for the study. Isocentre point dose measurements were carried out on a Varian iX linear accelerator using five commercial detectors in the CIRS Model 009 Cube 20 phantom. Measurements were made with IBA CC01 and CC04 compact ionisation chambers, IBA EFD^{3G} and PFD^{3G} diodes, and a PTW 60019 microDiamiond detector. Detector measurement repeatability was investigated and quantified by repeat measurements over three measurement sessions. The calculated dose was computed in both Pinnacle, using both 4 degree and 2 degree per control point gantry spacing (GS), and RayStation treatment planning systems. The agreement between measured and calculated dose was evaluated for each detector and calculation algorithm. A selection of established and novel aperture complexity metrics were calculated for the plan cohort. Correlations between complexity metrics and point dose discrepancy results were investigated.

Results: A statistically significant difference in mean measured dose between the CC04 chamber and all other detectors was found at the 95% confidence level. The between measurement sessions standard deviation was less than 0.5% of mean measured dose for all detectors excluding the PFD^{3G}. The CC01 achieved the greatest repeatability followed by the CC04, EFD^{3G}, microDiamond, and PFD^{3G}. A statistically significant difference in mean calculated dose was found for Pinnacle (both 4 and 2 degree GS) and RayStation calculation algorithms. For both 2 degree and 4 degree GS the mean point dose discrepancy is less than 0.55% for all detectors.

Statistically significant linear relationships were found, with weak to moderate strength Pearson correlation coefficients, for the following established complexity metrics MCSv, PI, PM, CAS, CLS, and MAD metrics. The strongest Pearson correlation coefficient, r = 0.407, was found for the PI metric with CC04 measured and Pinnacle 2 degree GS calculated dose. Evaluation of plan complexity in progressively smaller ROI centred on the dose measurement and calculation point increases the correlation strength for some complexity metrics.

Conclusion: Both the choice of a dose calculation algorithm and detector have a significant influence on point dose discrepancy results. Consequently, the strength of correlations between complexity metrics and point dose discrepancy is algorithm and detector specific. Therefore, the utility of individual complexity metrics to identify plans likely to fail QA will be department specific. The poor correlation strength of complexity metrics with point dose discrepancy results limits their clinical usefulness in identification of plans likely to fail QA.

Table of contents

List of figures

List of tables

Glossary

3DCRT	*Three-dimensional conformal radiation therapy*
ANOVA	*Analysis of variance*
CAS	*Cross axis score*
CLS	*Closed leaf score*
COR	*Coefficient of repeatability*
CP	*Control point*
CPE	*Charged particle equilibrium*
CT	*Computed tomography*
CVD	*Chemical vapour deposition*
DICOM-RT	*Digital imaging and communications in medicine file format for radiation therapy*
DMLC	*Dynamic MLC*
DMPO	*Direct machine parameter optimisation*
EBRT	*External beam radiation therapy*
IMAT	*Intensity modulated arc therapy*
IMRT	*Intensity modulated radiation therapy*
LINAC	*Linear accelerator*
LT	*Leaf Travel*
LTMCSv	*Leaf travel modulation complexity score*
MAD	*Mean aperture displacement*
MCSv	*Modulation complexity score for VMAT*
MFA	*Mean field area*
MLC	*Multi-leaf collimator*
MU	*Monitor unit*
MV	*Mega-voltage*
NTCP	*Normal tissue complication probability*
PI	*Plan irregularity*
POI	*Point of interest*

PM	*Plan modulation*
PMMA	*Poly(methyl methacrylate), perspex*
PMU	*Plan normalised monitor units*
PTV	*Planning target volume*
QA	*Quality assurance*
ROF	*Relative output factor*
ROI	*Region of interest*
SAS10	*Small aperture score with 10 mm threshold*
SAS20	*Small aperture score with 20 mm threshold*
SCDD	*Synthetic single crystal diamond detector*
SSD	*Skin surface distance*
TCP	*Tumour control probability*
TERMA	*Total energy released per unit mass*
TI	*Therapeutic index*
TMR	*Tissue maximum ratio*
TPS	*Treatment planning system*
VMAT	*Volumetrically modulated arc therapy*

Chapter 1

Introduction

Radiation therapy is an effective treatment for cancer when used alone or in conjunction with other treatments such as chemotherapy and surgery. It is suggested that 52% of cancer patients should receive radiation therapy at some stage of their treatment [1]. With new scanning techniques allowing for earlier diagnosis it is expected that radiation therapy will become more popular for the treatment of these localised malignancies [2]. The medical community must therefore address this issue by maximising the efficiency of cancer therapy in regards to time and cost.

1.1 Radiation Therapy

The general aim of radiation therapy is the eradication of cancer cells while sparing normal tissue, to avoid complications (morbidity). The principle is best illustrated by the plotting of tumour control probability (TCP) and normal tissue complication probability (NTCP) curves as a function of dose, as displayed in figure 1.1. Absorbed dose is the mean energy imparted per unit mass by ionising radiation and has units of gray (Gy), 1 Gy $= 1$ J kg^{-1}. The optimal radiation therapy dose for a given treatment will maximise TCP while minimising NTCP. The

figure shows an ideal case, with a large separation between the two curves allowing high TCP for the clinically acceptable NTCP of 5%.

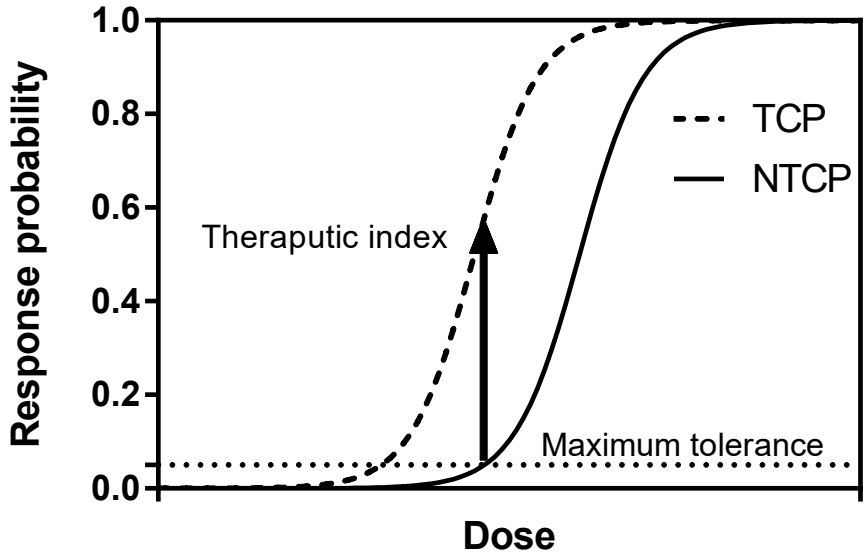


Fig. 1.1 Sigmoidally shaped dose response curves for tumour control probability (TCP) and normal tissue complication probability (NTCP). The therapeutic index (TI) is the ratio of expected TCP to NTCP at a clinically mandated maximum tolerance (5% in this example).

The shape of the response curve is dependent on the radiobiology of both the tumour cells and the normal tissue cells, and will depend on the treatment site. In some cases the TCP curve is shallower than the NTCP curve, due to the highly heterogeneous nature of malignant tissues, leading to a reduction in therapeutic index [3]. Modern radiation therapy allows the maximisation of the therapeutic index through the use of highly conformal treatments. Dose to malignant tissue is maximised while normal tissue is spared as much as possible, this can be visualised as a shift of the TCP curve to the left.

1.1.1 Linear Accelerators

Modern external beam radiation therapy (EBRT) is most commonly delivered by mega-voltage (MV) medical linear accelerator (linac) systems. Medical linacs can produce both electron and

photon beams for radiation treatment. The isocentre is defined as the point in space in which the linac gantry, collimators, and treatment couch rotate. The path through the waveguide and linac head for a treatment beam can be seen in figure 1.2. Within a linac, electrons are

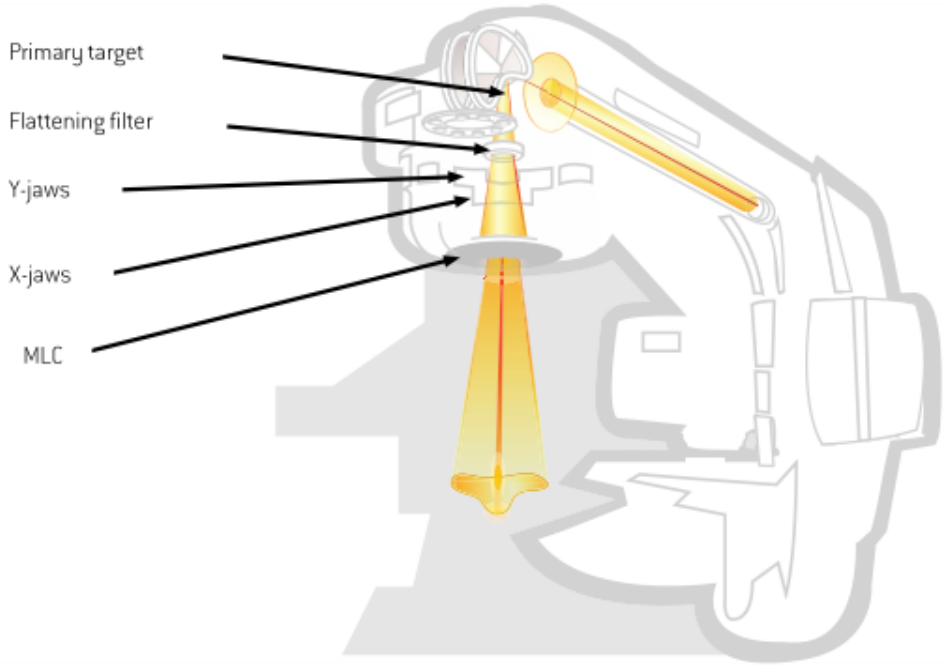


Fig. 1.2 Diagram of linear accelerator head components, reproduced from RayStaion 5 Reference Manual.

accelerated cyclically along a straight line path in the accelerating waveguide to energies of 6 - 18 MeV. Through the use of electric and magnetic fields, the beam transport system steers and focuses the electrons into a narrow and approximately mono-energetic beam. The electron beam is directed onto the primary target and produces a spectrum of Bremsstrahlung photons, with mean energy approximately one third of the maximum incident electron energy. The primary target can be thought of as a point photon source. The photon fluence generated from the primary photon source is then attenuated by a flattening filter to produce a beam with flat profile. The beam is then collimated to the desired shape via the combination of a primary collimator, moveable collimator jaws, and finally through a multi-leaf-collimator (MLC). The collimator jaws consist of thick tungsten blocks which define the maximum rectangular extent

of the treatment field.

The MLC consists of 2 banks of opposing tungsten leaves which can be retracted or extended into the beam. Neighbouring MLC leaves can move independently to define complex shapes, allowing the beam to conform to the target in two dimensions. Figure 1.3 shows a Varian Millennium 120 leaf MLC (Varian Medical Systems, Palo Alto, USA).

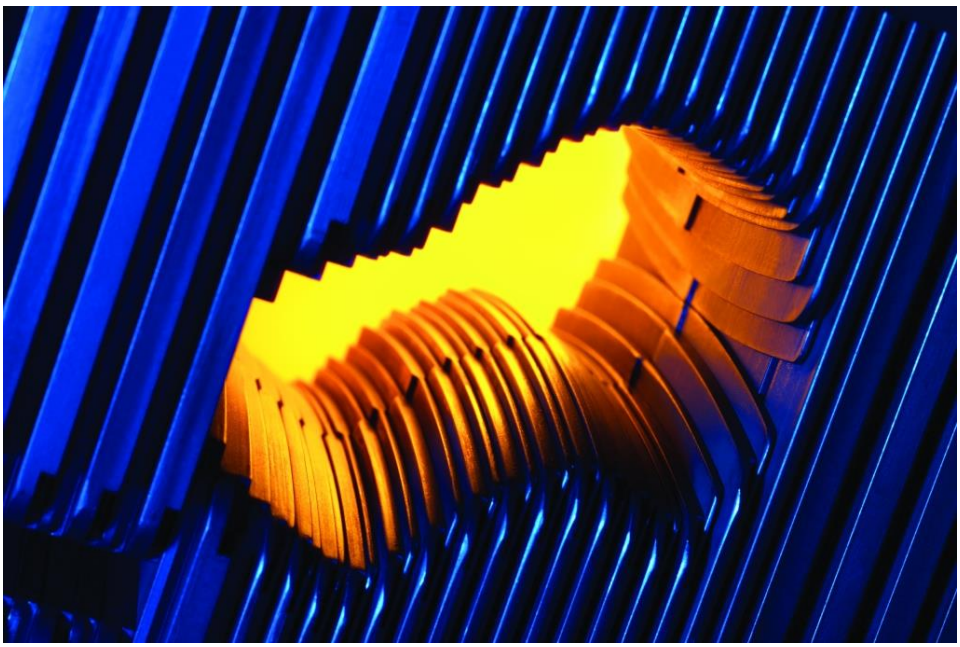


Fig. 1.3 Varian Millennium 120 leaf MLC, reproduced from Varian.

MLC leaf width is defined by their projected width at isocentre, ranging from 5 mm to 10 mm for non-stereotactic systems. To ensure constant attenuation with off-central-axis position, MLC leaves used in Varian linear accelerators have rounded leaf tips. The MLC does not provide perfect collimation of the primary beam, there will be some degree of inter- and intraleaf transmission. Tongue-and-groove construction is used to reduce interleaf transmission in the Varian system. Linac output is expressed in monitor units (MU), the calibration procedure is explained further in section 2.5.2.

1.1.2 Dose Calculation

Three major interaction processes occur within the patient for therapeutic photon beams: the photoelectric effect, Compton scattering, and pair production [4]. These interactions lead to attenuation and scattering of the incident photon beam within the patient. The photon radiation used in EBRT is not directly ionising. The interaction of incident photons with matter causes the liberation of directly ionising particles (electrons). It is these directly ionising particles which transfer energy along their ionisation track, and therefore dose, to the medium.

The calculation of dose and planning of radiation therapy treatments is undertaken electronically using a treatment planning system (TPS). Modern TPS dose calculation algorithms are model-based, where measured beam data is used to characterise the beam attributes. Dose profiles are measured in water at various depths for a selection of field sizes during commissioning of the linac and TPS. Figure 1.4 shows a series of cross axis field profiles and depth dose profiles for a 6 MV photon beam (6X), the measured dose profiles vary with beam energy, field size. and measurement depth [5].

The process of adjusting beam model parameters to increase the agreement between measured dose and calculated dose is known as beam modelling. These beam model parameters include the size and position of the primary source, the energy spectrum of the photon source, and parameters specific to the collimation system. Once the sufficient level of agreement between measured and calculated dose is achieved the measured data is no longer used. This allows the accurate calculation of dose for arbitrary beam configurations and geometry. The effect of beam modifiers, patient surface, and tissue inhomogeneity on dose distributions can be accounted for using a model based algorithm.

The first step of dose calculation is the computation of the energy fluence distribution from the linac. The computation considers the path of photon radiation from the primary source through the linac and accounts for the influence of the flattening filter and collimation systems which were characterised in the beam modelling process. A secondary source is used to model the

scattered radiation produced within the accelerator head. Model parameters used to characterise the linac differ between treatment planning systems, and depend on the implicit assumptions and simplifications of each algorithm.

Both treatment planning systems in this study make use of a collapsed cone convolution superposition algorithm for final dose calculation. The following section will provide a basic overview of the algorithm; refer to following references for a more in-depth description [6–9]. The algorithm separates the primary photon transport and the secondary transport of the photons and electrons generated by the primary photon interactions. The most common primary interaction for MV photon beams in tissue is Compton scattering, producing a free electron and a lower energy photon. The free electron deposits energy via ionisation within a short range of the primary interaction. Pre-calculated point spread kernels are used to characterise the spatial energy distribution resulting from each primary interaction. For a homogeneous medium and non-divergent beam, the dose at a specific point, r, is given by:

$$D(r) = \int \frac{\mu(r')}{\rho_m(r')} \Psi(r') A(r - r') dr' \tag{1.1}$$

where $\frac{\mu(r')}{\rho_m(r')}$ is the linear mass attenuation coefficient at position r', $\Psi(r')$ is the photon energy fluence, and $A(r - r')$ is the point spread kernal.

The product of linear mass attenuation coefficients and photon energy fluence, $\frac{\mu(r')}{\rho_m(r')}\Psi(r')$, gives the total energy per mass unit (TERMA). Therefore the dose distribution is given by the convolution of primary energy deposits (TERMA) with secondary energy deposits in the volume surrounding the primary interaction, characterised by the point spread kernel.

The algorithms used in both commercial treatment planning systems investigated account for tissue inhomogeneities and the divergent nature of the beam. TERMA is computed by the projection of the incident poly-energetic energy fluence through a density representation (computed tomography scan) of the patient. Dose is then computed via 3-D superposition of the TERMA with density scaled point spread kernels to account for tissue heterogeneities. A more

in-depth discussion of each algorithm can be found in the Pinnacle 9.10 Physics Reference Guide and the RayStation 5 Reference Manual [10, 11].

1.1.3 Intensity Modulated Radiation Therapy

Traditionally external beam photon radiation therapy consisted of the delivery of treatment beams with uniform intensity across the field. The process of selecting a suitable beam arrangement to produce the prescribed dose distribution is known as forward treatment planing. Wedges and compensators were used to modify the intensity profile of these beams to achieve a desired dose distribution within the target. While these simple techniques do modulate the intensity of the treatment field, they are not included in what has become known as intensity modulated radiation therapy (IMRT). The principle of IMRT is to modulate the intensity of incident radiation in order to achieve a higher degree of conformity for the resulting dose distribution and the target volume. The theoretical framework of the technique can be traced back to 1982 when Brahme proposed what is now known as the definition of the inverse planning problem in IMRT: "What is the desired lateral dose profile in the incident beam that produces a desired absorbed dose distribution in the body?" [12]. The beam modulation is optimised to increase the intensity of incident radiation passing through the target volume while also reducing the intensity of incident radiation travelling through normal tissue. This results in both increased target volume dose and reduced dose to critical structures around the target volume compared to conventional radiation therapy [13]. The improved target dose conformity results in better locoregional control, while the increased sparing of normal tissue leads to the reduction of both acute and late toxicity [14].

In modern IMRT, beam modulation is achieved through either step-and-shoot or dynamic MLC (DMLC) techniques. In step-and-shoot IMRT, the beam is halted while MLC leaves move to designated positions then remain stationary while the beam is on; this process is repeated for all segments in each beam. In DMLC delivery the MLC leaves move continuously, at variable

speeds, while the beam is on, reducing treatment time when compared with step-and-shoot IMRT. In both step-and-shoot and DMLC IMRT the gantry angle is static for each beam. The desired final dose distribution results from the superposition of dose from each individual beam segment. Often over 100 segments are defined across all beams in a treatment plan.

Traditional forward planning methods are unsuitable for the generation of IMRT, leading to the development of inverse planning first proposed by Brahme. The planner will specify clinical objectives for target volumes, often including maximum and minimum dose and dose to volume constraints. Similarly for healthy organs and critical structures, the planner will specify limiting dose constraints. An inverse planning optimisation algorithm can be broken down into two main components: an objective function which encapsulates the clinical objectives and a method to maximise/minimise the objective function. The inverse planning algorithm will then analytically, or through iterative methods, optimise intensity profiles in order to generate a dose distribution which fulfils the clinical objectives. The optimisation process is viewed as a "black box" and therefore verification of the resultant treatment plan is required, as discussed further in section 1.2.

1.1.4 Volumetric-Modulated Arc Therapy

Intensity modulated arc therapy (IMAT) was first proposed in 1995 [15]. Yu described and demonstrated feasibility for this technique in which the gantry is rotated around the patient while delivering radiation and the overlapping of multiple arcs provided intensity modulation. Until relatively recently implementation of the technique was limited, with linac control systems and inverse treatment planning tools not being sufficiently robust for clinical use.

Now both Elekta (Elekta, Stockholm, Sweden) and Varian offer linacs capable of delivering IMAT. Commercial availability along with improvements in inverse treatment planning systems and increasing computational power make the deliverability of IMAT clinically feasible. Volumetric modulated arc therapy (VMAT) is an extension of IMAT to include dynamic variation of

dose rate, gantry rotation speed, and MLC positions to deliver highly conformal dose to target structures [16]. The variation of these extra parameters allows treatment to be delivered over a fewer number of arcs than IMAT, greatly improving the time efficiency of the treatment. The reduction in treatment time and MU, when compared to conventional IMRT, also leads to a reduction in integral radiation dose to the rest of the body by reducing linac head leakage [17]. Like conventional IMRT, inverse planning is used to generate VMAT plans. The optimisation and dose calculation algorithms for the two treatment planning systems used at Auckland City Hospital are discussed in section 2.2.

1.2 Patient Specific Quality Assurance

Due to the complex nature of IMRT and VMAT treatment delivery and dose calculation when compared to traditional 3D conformal radiation therapy there is increased uncertainty in the agreement between the TPS calculated dose and delivered dose for IMRT and VMAT. IMRT plans consist of complex arrangements of many small and irregular MLC field segments. Therefore, accurate MLC leaf positioning (and leaf speed for DMLC) during delivery must be assured. VMAT represents an even more complicated scenario by the inclusion of dynamic gantry rotation and dose rate variation. Along with concerns of treatment deliverability, accurate dose calculation for VMAT is challenging [18]. Therefore pre-treatment patient specific quality assurance (QA) is performed to verify the agreement between TPS calculated dose and delivered dose measured in phantoms.

The QA process typically involves the calculation of dose for the patient specific treatment plan on a CT scanned phantom; calculated dose is then compared with measurements made on the phantom. The AAPM Task Group 120 report provides recommendations for the use of dosimeters, phantoms, and dose distribution analysis techniques during IMRT/VMAT quality assurance [19]. Two measurement approaches are commonly used for dosimetric evaluation

of IMRT/VMAT treatment plans: two-dimensional dose plane measurements and point dose measurements.

1.2.1 Planar Dose Analysis

Two-dimensional dose plane measurements, using film or chamber arrays, provide dosimetric and spatial information of the delivered dose distribution. The agreement between dose distributions can be characterised by both the dose-difference at each dose grid point and the distance-to-agreement, which is a measurement of the distance discrepancy between two dose distributions. Gamma analysis combines both dose-difference and distance-to-agreement criteria into a single index; the percentage of dose grid points meeting the specified criteria [20, 21]. Auckland City Hospital requires at least 95% of points meet the (3% / 3 mm) dose-difference and distance-to-agreement criteria for absolute gamma; the pass rate is relaxed to 85% for the stricter (2% / 2 mm) criteria.

Limitations of gamma analysis alone for patient specific QA have been summarised in the literature [22]. The sensitivity of gamma analysis in detecting dosimetric errors has been shown to strongly depend on the choice of analysis criteria. For example, it has been shown that the (3% / 3 mm) criteria is insufficient to detect some introduced systematic errors [23, 24].

1.2.2 Point Dose Analysis

Point dose measurements provide a measurement of absolute dose at a pre-determined point of interest (POI), typically within the target volume. In point dose measurements a detector is placed within a phantom to measured absolute dose at the POI, which is then compared with the TPS calculated dose at the POI. The discrepancy between the measured and TPS calculated dose is subsequently evaluated; at Auckland City Hospital agreement must be within 3% for all clinical IMRT and VMAT plans.

Kry et. al evaluated the sensitivity of point dose measurements and planar dose analysis (using

both film and diode arrays), in the detection of dosimetric errors in IMRT delivery [25]. The combination of both point dose measurement and planar analysis achieved sensitivity of 54%, point dose measurements alone resulting in 25%, and planar analysis alone using a diode array only 14%. These results highlight the importance of supplementing planar dose measurements with point dose measurements to maximise the ability to detect dosimetric errors.

1.2.3 Plan Complexity Metrics

The degree of beam modulation can be considered analogous to the complexity of the control points, and indeed the entire treatment plan. Various studies have outlined the disadvantages associated with highly complex treatment plans; these include increased beam on time leading to increased risk of patient movement, increased mechanical stress on the treatment unit, and higher probability of the actual delivered dose deviating from the planned dose [26, 27].
Patient specific QA is time consuming and serious concerns have been raised over the sensitivity of existing measurement methods in detecting clinically relevant errors. The development of a method to identify treatment plans likely to fail QA would allow more efficient use of the limited machine and physicist time. Many complexity metrics have been proposed in order to characterise treatment plans and identify those likely to fail QA. Two general approaches have emerged: direct analysis of beam aperture or analysis of the resulting plan intensity fluence maps. Aperture based metrics represent the most direct method to investigate the influence of key plan parameters on both TPS calculation and treatment machine deliverability, and are therefore investigated in this study.

1.3 Radiation Dosimeters

In order to measure absorbed dose in a medium, the introduction of a radiation sensitive device into the medium is required. Generally the sensitive volume of the detector will be constructed

of a material different from the medium of interest. Cavity theory has been developed to relate the absorbed dose measured within the detector's sensitive volume to the absorbed dose in the medium. Spencer-Attix-Nahum cavity theory is an extension of work by Louis Harold Gray and William Lawrence Bragg, and accounts for deviations from the idealised formulation of Bragg-Gray cavity theory [28].

Bragg-Gray cavity theory describes an idealised situation where the following conditions are met:

1. The volume of the cavity is small when compared to the range of incident charged particles as to not perturb the fluence of charged particles in the medium;
2. Absorbed dose in the cavity is only deposited by charged particles crossing the cavity.

The first condition implies the electron fluence within the cavity is equal to the equilibrium fluence in the surrounding medium. This condition is only met in regions of charged particle equilibrium (CPE), where the number of charged particles leaving a volume is equal to the number entering. While the second condition rules out the production of secondary electrons within the cavity, all dose is due to electrons crossing the cavity and no electrons stop within the cavity. Assuming these conditions are true then dose to the medium, D_{med}, is given by:

$$D_{med} = D_{cav} \left(\frac{\overline{S}}{\rho} \right)_{med,cav} \tag{1.2}$$

where D_{cav}, is the dose to the cavity, and $\left(\frac{\overline{S}}{\rho} \right)_{med,cav}$ is the ratio of the mean unrestricted mass collision stopping powers of the medium and the cavity.

Spencer-Attix-Nahum cavity theory extends Bragg-Gray cavity theory to account for secondary electrons (delta rays) produced by hard collisions between the primary electrons and cavity material. This results in modifications to the formulation of stopping power ratios, and increased agreement with experimental results when compared to Bragg-Gray Cavity theory [29, 30].

Modern dosimetry protocols, including IAEA TRS-398 (followed at Auckland City Hospital),

are based on the Spencer-Attix-Nahum formulation which is applicable to megavoltage photon and electron beams [31–33]. Detector response in standard reference conditions (10×10 cm^2 field at 10 cm depth) is well characterised by these dosimetry protocols due to the presence of CPE or transient CPE, thereby meeting the Bragg-Gray conditions. But VMAT plans often consist small and irregular fields, which differ greatly from standard open field reference conditions. Lateral electronic disequilibrium is observed in the small fields often found in VMAT treatments [34, 35]. VMAT treatment plans often require steep dose gradients to achieve the desired target dose while sparing organs at risk. These conditions pose a challenge for point dose measurements as volume averaging effects are significant in high dose gradients (especially for ionisation chambers). The non-uniformity of response within the detector volume and perturbation factors should also be considered [36].

It is for these reasons dose measurements in VMAT fields are challenging, in this study the suitability of five detectors for VMAT point dose measurements have been investigated. The five detectors fall into three main classes of detector; ionisation chambers, silicon diodes, and synthetic single crystal diamond detectors.

1.3.1 Ionisation Chambers

Energy from an incident ionising radiation beam is deposited by the production of ion pairs within the medium. The basic construction of an ionisation chamber consists of a gas filled cavity surrounded by a conductive outer wall and central charge collecting electrode. A polarising voltage is applied across outer wall and central electrode to ensure ion pairs are collected. Polarising voltages are of the order of a few hundred volts to ensure efficient collection (typically between 95% and 100%) but avoid secondary ion pair production by the motion of primary ion pairs [37]. Leakage currents are reduced by the addition of a guard electrode. The mean dose deposited, D, within the mass of air, m, enclosed by the chamber is

related by the charge collected, Q, by

$$Q = \frac{emDP_{ion}}{W} \tag{1.3}$$

where e is the elementary charge, P_{ion} is a correction factor for the recombination of ions within the charge collection volume, and W is the mean ionisation energy in air. Therefore, dose can be determined by the measurement of charge using an electrometer connected to the ionisation chamber.

1.3.2 Silicon Diodes

Silicon diodes used in radiation dosimetry consist of a disc of p-type silicon counter-doped with a thin layer of n-type silicon, to form a p-n junction diode. At the interface of the n- and p-type material diffusion of majority carriers (electrons from n-region leaving positively charged donor ions, and holes from the p-region leaving negatively charged acceptor ions) across the junction takes place, forming the depletion region. Equilibrium is reached when the electric field created across the depletion region by the donor ions opposes further flow of majority carriers. Electron-hole pairs are created during irradiation within the body of the dosimeter (including the depletion zone). Minority carriers produced within a diffusion length diffuse into the depletion region, otherwise recombination occurs [38]. The length of the sensitive region of the detector is therefore the depletion region length plus the diffusion length of the minority carriers. The intrinsic potential across the depletion region makes charge collection possible without the requirement of external bias voltage. See figure 1.5 for schematic of the use of a p-n junction diode as a radiation detector. Charge carriers are accelerated across the depletion region by the electric field resulting from the intrinsic potential. There exists a linear relationship between the measured charge and dose when the diode is operated in short circuit mode.

The relative efficiency per unit volume for silicon diodes is on the order of 10^4 times greater than an air ionisation chamber [39]. This is the result of two factors: Si is approximately 2000 times denser than air, and the mean ionisation energy for Si is approximately a tenth of that in air (around 3 eV). This allows the construction of Si diode dosimeters with much smaller sensitive volumes when compared with ionisation chambers.

1.3.3 Synthetic Single Crystal Diamond Detector

The use of diamond detectors in small field dosimetry is well established [40]. Diamond detectors achieve spatial resolution comparable with silicon diodes, but provide superior tissue equivalence and radiation hardness (resistance to radiation damage). The band gap for natural diamond is approximately 5 times larger than silicon (5.54 eV versus 1.12 eV for silcon), resulting in lower leakage currents due to fewer free charge carriers at room temperature. When diamond is irradiated, a temporary change in electrical conductivity is observed resulting from the production of electron and hole pairs which have sufficient energy to traverse the crystal. The use of natural diamond crystals for dosimetry is impacted by low availability and reproducibility of 'detector grade' crystals and the high costs associated.

These challenges motivated the development of synthetic single crystal diamond detectors (SCDDs) produced by chemical vapour deposition (CVD) by Marinelli et. al at the Industrial Engineering Department of Rome Tor Vergata University in Italy [41]. The detectors consist of a multi-layered metal/nominally intrinsic/p-type structure, as seen in figure 1.6. It was found that the if barrier junction is located between the Al contact and the nominally intrinsic diamond layer, the detector acts as a Schottky barrier photodiode which can be operated with both zero bias voltage (photovoltaic mode) or with a positive voltage applied to the Al contact. When the detector is irradiated in photovoltaic mode the total signal consists of electron/hole pairs generated in the depletion region and charges generated in the neutral zone that have diffused into the depletion region.

1.4 Objectives

The patient specific quality assurance requirements for VMAT are greater than for traditional 3D-conformal treatments. An increasing proportion of patients are being treated with VMAT at Auckland City Hospital, including palliative treatments. The increase in physics workload spent on QA of these VMAT treatments is unsustainable, therefore the VMAT QA process must be optimised. This requires a deeper understanding of the relationship between calculated dose and measured dose in terms of plan complexity.

The main goal of this project will be to identify and develop an understanding of the salient parameters influencing plan complexity and the impact of these on the accuracy of point dose measurements in VMAT fields. The choice of detector for the point dose measurement will also be explored. The research will allow especially complex treatment plans to be identified during planning and potentially influence the patient specific quality assurance undertaken.

Point dose measurements will be remeasured for a cohort of past clinical treatment plans. The influence of treatment planning system dose calculation on point dose results will be investigated. Detector choice will be investigated concurrently by comparing results for various detector types; small volume ionisation chambers, diodes, and synthetic diamond detectors will be compared. Analysis of both repeatability and agreement with treatment planning system calculation will be undertaken for each detector.

Software to read and extract essential information from DICOM-RT files generated by the treatment planning system will be developed. Existing complexity metrics proposed in the literature will then be calculated. Analysis of these against point dose measurement results will allow key parameters to be identified and will steer development of additional complexity metrics. Finally, this study will determine whether complexity metrics can be used to reliably predict quality assurance results.

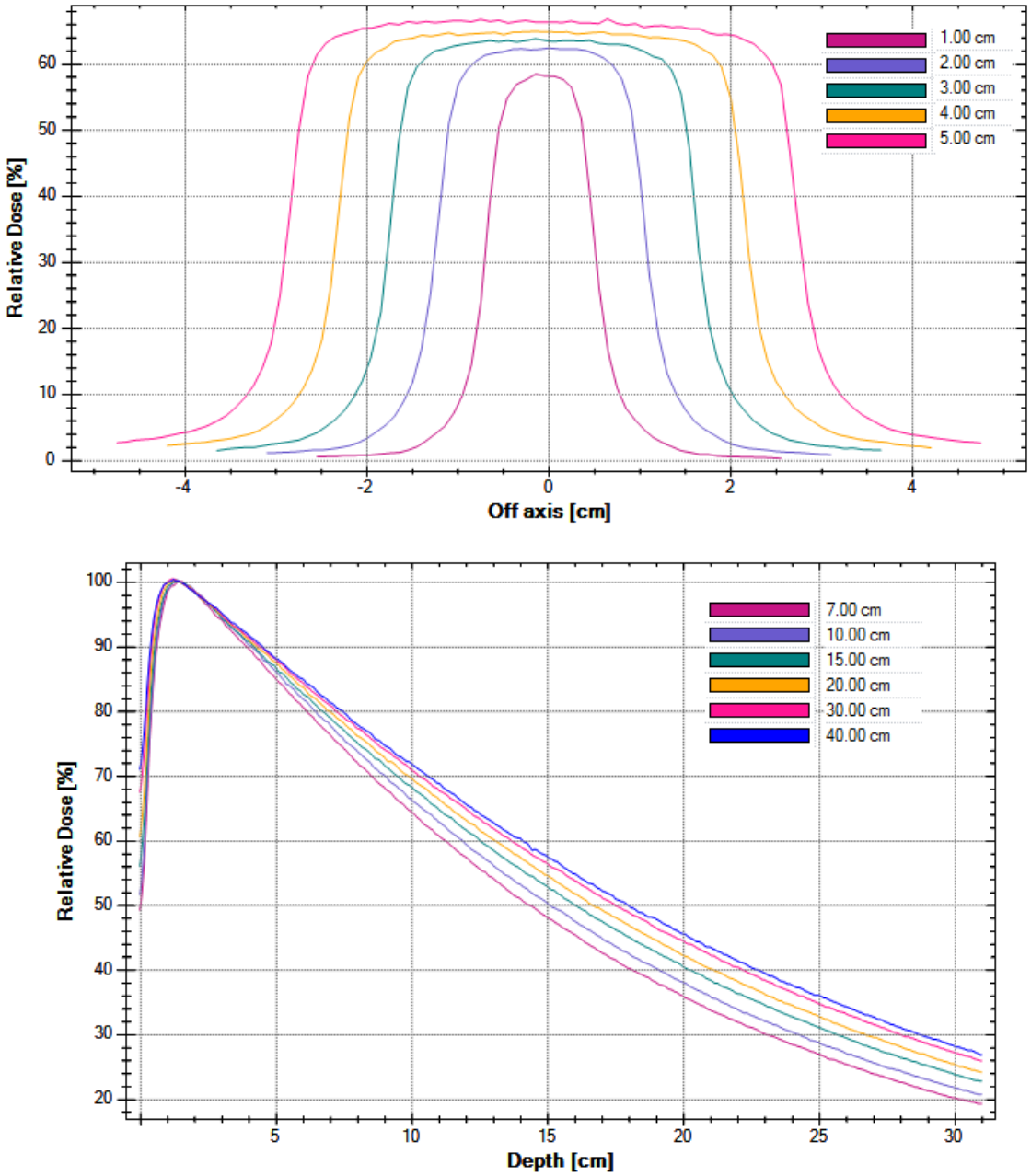


Fig. 1.4 Example beam dose profile measurements used in TPS commissioning of a 6X beam, square field size given in legend. Top: measurements of dose profile across the treatment field at 10 cm depth. Bottom: the variation of dose with depth, normalised to dose maximum.

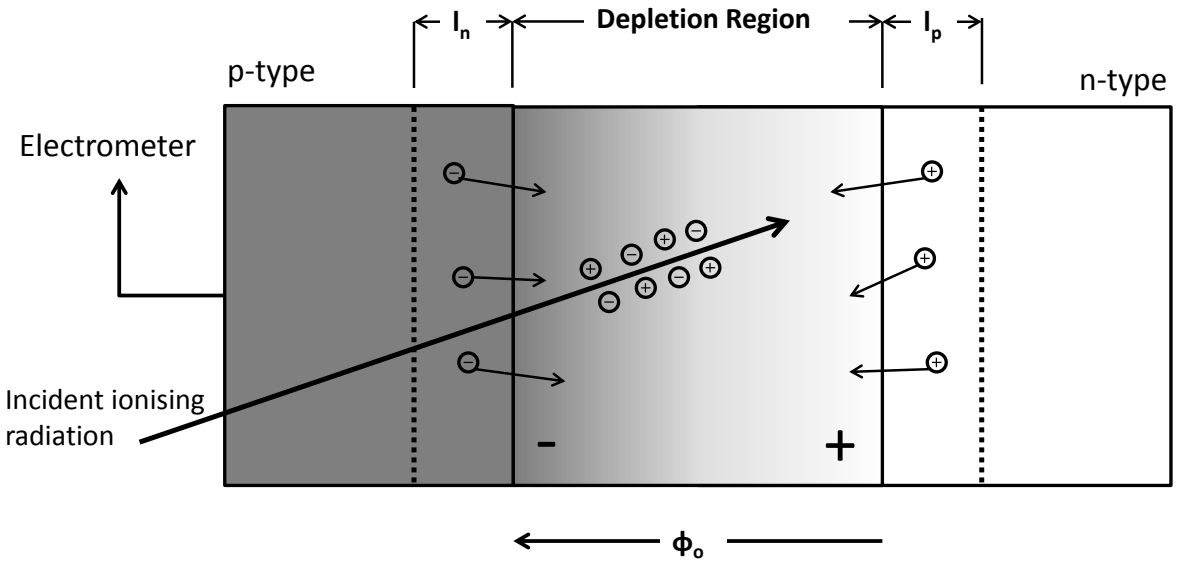


Fig. 1.5 Schematic of a silicon diode dosimeter. Incident ionising radation produces electron-hole pairs in the body of the dosimeter and within the depletion region. Excess minority carriers created in the body of the dosimeter within one diffusion length, l_n and l_p, for p-type and n-type respectively, diffuse into the depletion region. Carriers are swept across junction by intrinsic potential, ϕ_o, net current flows in reverse direction of diode.

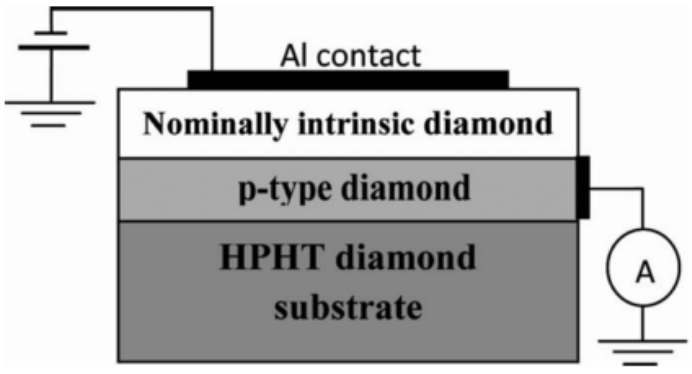


Fig. 1.6 Schematic of the synthetic single crystal diamond detector. Reproduced from Almaviva et al. 2010 [42]

Chapter 2

Methods and Materials

2.1 Treatment Plans

All VMAT plans investigated in this study were planned and delivered clinically at Auckland City Hospital between April and June 2015. The plans were generated using the SmartArc module in Pinnacle^3 (Version 9.8 and 9.10, Philips Medical, Fitchburg WI) following departmental planning protocols, using 6 MV photons. Treatment doses were calculated using the collapsed cone convolution superposition algorithm. The plans were optimised using 4 degree gantry angle spacing between control points with a 3 mm resolution dose grid. All plans were delivered using Varian iX linear accelerators with Millennium 120-leaf MLC (Varian Medical Systems, Palo Alto, USA).

It has been shown that treatment plan complexity correlates strongly with target site and shape [43]. In order to isolate the dosimetric impact of overly modulated treatment plans the decision was made to limit this study to a single treatment site. A cohort of 45 prostate single arc 6X VMAT plans were selected at random. Two standard prescriptions were present in the 45 plans. 23 with a prescription of 74 Gy to prostate PTV with 65 Gy to seminal vesicle PTV in 37 fractions, and 78 Gy to prostate PTV in 39 fractions for the remaining plans.

2.2 TPS Dose Calculation

Two commercial treatment planning systems are commissioned for clinical external beam radiation therapy at Auckland City Hospital; Pinnacle and RayStation (Version 4.5, RaySearch Laboratories, Stockholm, Sweden).

2.2.1 Pinnacle

The SmartArc module makes use of an optimisation algorithm developed by Bzdusek et al [44]. The optimisation process begins with the generation of a coarse set of segments spaced in 24 degree intervals within the user defined arc. Intensity modulation optimisation is carried out for each segment and resulting fluence maps are converted to sliding window MLC segments. The MLC segments are filtered, redistributed, and finally interpolated across the arc to achieve evenly spaced 8 degree CP spacing. These control points are used in the Direct Machine Parameter Optimisation (DMPO), which is applied to leaf positions, MU, dose-rate, and delivery time for each control point. The optimisation is constrained by the machine specifications, such as maximum gantry speed, MLC leaf speed, and leaf travel per degree. Another linear interpolation is carried out to generate a second set of control points to achieve the user defined angular spacing. These control points are not part of the DMPO and are only used to increase dose computation accuracy, they must be regenerated every time dose is computed. A fast pencil beam algorithm, based on singular value decomposition, is used for dose calculation during optimisation to keep computational time clinically acceptable. After machine parameters have been optimised, the collimator jaws are set to the extent of the largest open leaf positions across all segments.

Final dose calculation is performed using the collapsed cone convolution superposition algorithm at user defined control points. Various authors have postulated that an increase in dosimetric agreement can be achieved by reducing the gantry angle sampling interval for final dose calculation [45]. This comes at the expense of longer calculation times. For clinical plans,

Auckland City Hospital currently uses 4 degree gantry angle spacing between control points as a compromise between calculation accuracy and speed. For final dose calculation in this study, the plans were recalculated with using interpolated 2 degree gantry angle sampling, so that the 4 degree and 2 degree calculations could be compared.

2.2.2 RayStation

RayStation differs from Pinnacle in that dose is not calculated at each control point for VMAT treatment plans. Instead, the dose calculation is split into arc segments which represent the region between two control points. These arc segments are further divided into sub-segments of equal angle to ensure maximum MLC leaf travel within each sub-segment is less than 1.5 cm. The monitor units of the arc segment are distributed over sub-segments and MLC leaf positions are interpolated to the centre position of each sub-segment. The fluence of each sub-segment is calculated and summed to obtain the arc segment fluence, which is then applied to the centre position of the arc segment for dose calculation. TERMA is calculated by tracing the arc segment fluence into the patient at the mean gantry angle of the segment. In order to decrease computation time, TERMA is then accumulated over 10 degree gantry angle intervals. Sphere point spread kernel tracing is performed at the middle of each accumulated region.
The RayStation beam model was not commissioned using the same beam measurement data as the Pinnacle beam model. Therefore, RayStation provides an independent dose calculation method to compare with doses calculated in Pinnacle. All treatment plans in this study were recalculated in RayStation to investigate the influence of dose calculation algorithm on the correlation of complexity metrics and point dose measurement results.

2.3 Quantification of Plan Complexity

In-house software was developed in MATLAB (Version R2014a, MathWorks Inc., Natick, MA) to allow batch analysis of treatment plans exported from the planning system in DICOM-RT format. The software extracted plan parameters, including MLC leaf positions, jaw sizes, gantry angle, MU and dose-rate, for each control point. Using these parameters both established and novel plan complexity metrics were calculated for each plan.

2.3.1 Established Metrics

Table 2.1 Summary of initialisms of investigated established metrics.

MCSv	Modulation complexity score for VMAT
LT	Leaf travel
LTMCSv	Leaf travel modulation complexity score
PI	Plan irregularity
PM	Plan modulation
CAS	Cross axis score
CLS	Closed leaf score
MAD	Mean aperture displacement
MFA	Mean field area
SASx	Small aperture score with x mm threshold
PMU	Plan normalised monitor units

In order to characterise plan deliverability MLC leaf travel (LT) metrics were investigated. McNiven et al. first proposed the Modulation Complexity Score (MCS) for step-and-shoot IMRT, this was later adapted for VMAT by Masi et al. [45, 46]. The MCS provides an indication of the mobility of MLC leaves during delivery by considering two parameters: the aperture area variability (AAV), the variation in control point area relative to the maximum aperture defined by all control points; and the leaf sequence variability (LSV), the variation in field shape between control points, considering the change in leaf position between adjacent MLC leaves. All control points are weighted according to fractional MU of the beam. The MCS decreases from unity (defined as single rectangular field) to zero with the addition of

control points with smaller and more irregular apertures compared to maximum beam aperture. Du et al. have developed a set of metrics to quantify different aspects of plan complexity including plan aperture area (PA), aperture shape irregularity (PI), and plan modulation (PM) [47]. A simplified physical interpretation of these metrics follows: PA is the MU weighted average area of all MLC defined beam apertures in a plan, the PI describes the deviation of all segment aperture shapes from a circle, and the beam modulation characterises the degree to which the field is broken into small segments. Increased plan complexity can be characterised with decreased PA, and increased PI and PM.

Crow et al. introduced a set of metrics which quantify parameters most likely to contribute to inaccurate TPS dose calculation [48]. These include presence of small field segments, the proportion of apertures centred off axis, and the presence of closed MLC leaves within the jaw defined field. The cross axis score (CAS), closed leaf score (CLS), mean aperture displacement (MAD), mean field area (MFA), and small aperture score (SAS) metrics were implemented in this study.

2.3.2 Localised Complexity

Complexity metrics have historically been developed as a method of predicting gamma analysis results for measurements of the entire field fluence, such as with a detector array, thus they are calculated over the whole treatment field. The isocentre dose measurements in this study represent dose to a point. Therefore, various complexity metrics were calculated within a localised region about the point of interest with the goal to increase the sensitivity of the metric. Figure 2.1 illustrates the ROI definition and naming formalism for this study; a 1 cm ROI corresponds to a 1×1 cm^2 square centred on the isocentre. The influence of ROI size variation on metric correlation with point dose discrepancy has been investigated.

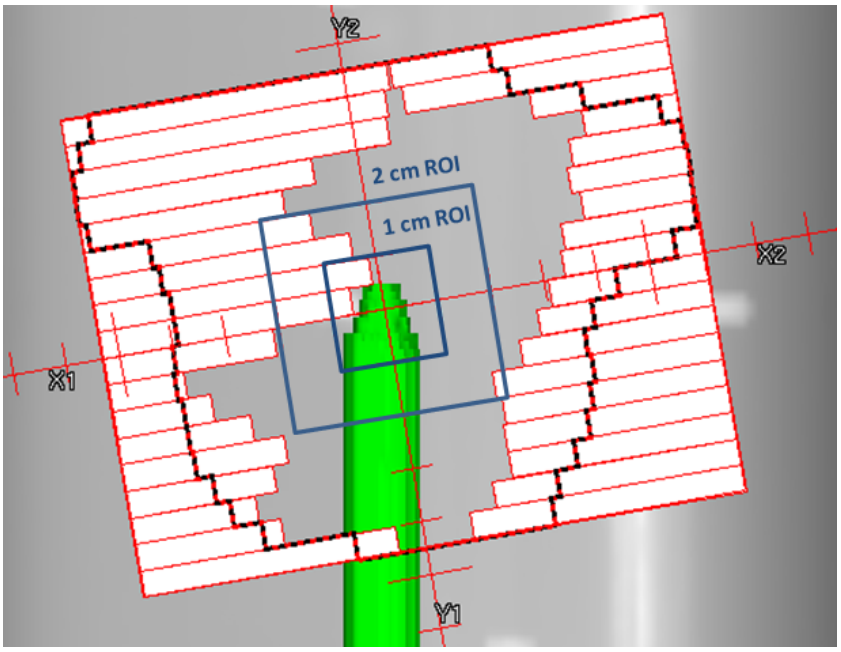


Fig. 2.1 Definition of ROI used for localised complexity metric analysis as seen from beam's eye view. MLC leaves are clipped at main collimator jaw positions. The green contour defines the air cavity for the CC04 CIRS Cube phantom insert. Both the 1 cm and 2 cm ROI are shown, corresponding to 1×1 cm^2 and 2×2 cm^2 squares centred on isocentre respectively.

2.3.3 Additional Metrics

The perimeter defined by MLC leaves was summed for each control point and averaged over the beam according to control point weight to give the mean field perimeter (MFP) metric. The ratio $\frac{\text{MFP}}{\text{MFA}}$ will increase for irregular field shapes where an increased proportion of field area is defined by MLC leaf edges.

The mean, variance, skewness, and kurtosis of the distribution of MLC leaf gap apertures was investigated. A low mean leaf aperture indicates the presence of small fields which are problematic for both the treatment dose calculation and the measurement. High variance and skewness indicate irregularly shaped fields. A high kurtosis indicates the presence of outliers in MLC aperture. These 4 distribution moments were also calculated for the dose rate distribution.

The blocked fraction (BF) is defined as the fractional area within an ROI about the isocentre

shielded with MLC leaves or collimator jaws. The BF varies from 0, where no MLC leaves are within the ROI, to unity in the case of a completely blocked region.

2.4 Dosimeters

For all measurements in this study, each detector was paired with the same PTW UNIDOS E T10009 (PTW-Frieburg, Germany) electrometer. Figure 2.2 shows the detectors investigated in this study.

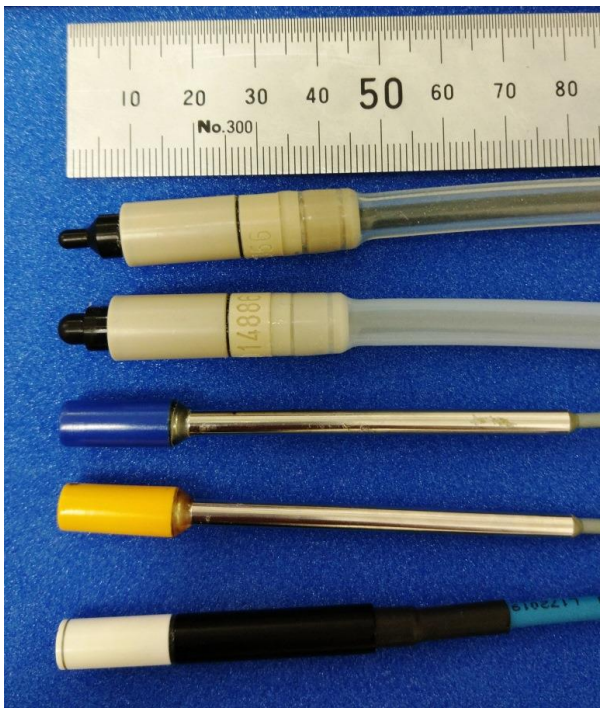


Fig. 2.2 Detectors investigated in this study, from top to bottom: CC01, CC04, EFD^{3G}, PFD^{3G}, and microDiamond.

2.4.1 Ionisation Chambers

Two ionisation chambers were investigated, the IBA CC01 and CC04 compact chambers (IBA Dosimetry GmbH, Schwarzenbruck, Germany). The active volumes of the chambers are 0.01 cm^3 and 0.04 cm^3 respectively. Both chambers have cavity length of 3.6 mm, with the CC01

and CC04 having outer electrode inner diameter of 2 mm and 4 mm respectively. The outer electrode for both chambers is constructed of Shonka (C-552) conductive plastic. See table 2.2 for a summary of the properties of each of the ionisation chambers.

Table 2.2 Summary properties of ionisation chambers used in this study.

Detector	Central electrode	Wall Thickness [mg/cm^2]	Approximate sensitivity [C/Gy]
IBA CC01	Steel	88	4×10^{-10}
IBA CC04	Shonka (C-552)	70	1.10×10^{-9}

2.4.2 Silicon Diodes

Two diodes have been used in this study, the IBA EFD^{3G} and PFD^{3G}. Both consist of a 2.1 $\times$ 2.1 $\times$ 0.4 mm chip, with active detector diameter and thickness of 1.6 and 0.08 mm respectively. Both diodes have a typical sensitivity of 100 nC/Gy. The PFD has an integrated energy filter that reduces the over-response of the diode to low-energy scattered radiation, which is abundant in large photon fields at depth. IBA states that the PFD is optimised for photon fields larger than 10 $\times$ 10 cm^2. IBA goes on to recommend the EFD for small photon fields less than 10 $\times$ 10 cm^2.

2.4.3 Synthetic Diamond

The PTW 60019 microDiamiond detector (PTW-Frieburg, Germany) is a commercialisation of the SCDD operating in the Schottky barrier photodiode configuration [49]. Figure 2.3 shows an X-ray image and geometrical sketch of microDiamond detector. The device consists of a 1 μm intrinsic diamond film sandwiched between a 3 x 3 x 0.3 mm^3 diamond plate and the 2.2 mm diameter Al contact, encased in a polystyrene and epoxy housing of 7 mm diameter. This results in an active surface area of 2.2 mm in diameter and 1.0 μm thick active volume

thickness, situated 1 mm below the top of the housing. For this study the detector was oriented perpendicular to the beam and operated in photovoltaic mode.

2.5 Point Dose Measurements

2.5.1 Phantom

Point dose measurements were made using a CIRS Model 009 Cube 20 phantom (CIRS, Norfolk, VA, USA), shown in figure 2.4. The phantom is constructed from Plastic Water DT, with physical density of 1.039 g cm^{-3} with an electron density of 3.345×10^{23} cm^{-3}, for 6 MV beams percent depth dose curves are within 0.5% of water [51]. Custom inserts were manufactured to position each detector at the centre of the phantom. They were manufactured in-house from Solid Water (Gammex rmi, Widdleton, WI, USA).

2.5.2 Preparation and Delivery

In order to compare calculated and delivered dose, each plan was recalculated on a CT image dataset of the CIRS phantom. In every plan the isocentre was within a high-dose region in the PTV, therefore the plan isocentre was set to the centre of the phantom. The CT couch was overridden and an ROI based model of the Varian Exact IGRT couch was applied. Dose calculation was performed with a 3 mm resolution dose grid using a collapsed cone convolution algorithm. The origin of the dose grid was selected such that a calculation point would coincide with the centre of the detector. The dose distribution was evaluated for each plan to ensure the measurement point did not correspond to a high dose gradient region. For all plans, dose heterogeneity across detector active volume was less than 5%, following AAPM TG 120 recommendations [19].

The treatment couch longitudinal and lateral positions were kept constant between measurement sessions to ensure accurate and repeatable results. The top surface of the phantom was posi-

tioned 90 cm from the photon source (90 cm SSD) using calibrated front pointers. Positioning of the phantom on the treatment couch was achieved by aligning the linac optical field with reference marks on the phantom. The custom inserts positioned the detector effective point of measurement along its central axis at the isocentre (10 cm deep).

Once the phantom and dosimeter were in position, a reference field was delivered to compensate for machine output fluctuations and allow conversion of raw chamber reading to absolute dose in Gy. The reference field was a 6 MV photon beam of 100 MU, with field-size of 4×4 cm^2 for the EFD and 10×10 cm^2 for all other detectors. The smaller reference field for EFD was chosen as unshielded diodes are known to over-respond to the low-energy scattered photons present in large fields. The reference field was delivered regularly throughout the coarse of each measurement session to ensure consistency of linac output. Each treatment plan was then delivered without interruption.

At Auckland City Hospital linacs are calibrated to deliver 1 cGy/MU to the isocentre at depth of dose maximum, d_{max}. Therefore for each plan, the ratio of raw electrometer reading to that of the reference beam was multiplied by a tissue maximum ratio (TMR) (and relative output factor (ROF) for the EFD) to calculate absolute dose at isocentre, see equations 2.1 and 2.2.

$$\text{Dose (Gy)} = \frac{M_{Plan}}{M_{Reference}} \text{TMR}(10 \text{ x } 10 \text{ cm}^2, 10 \text{ cm}) \tag{2.1}$$

$$\text{Dose}_{\text{EFD}}(\text{Gy}) = \frac{M_{Plan}}{M_{Reference}} \text{TMR}(10 \text{ x } 10 \text{ cm}^2, 10 \text{ cm}) \text{ ROF}(4 \text{ x } 4 \text{ cm}^2, 10 \text{ cm}) \tag{2.2}$$

2.5.3 Point Dose Repeatability

In order to quantify the repeatability of point dose measurements, 10 plans were selected to be remeasured in two additional measurement sessions. Each plan was measured once per measurement session, for a total of three measurements per plan per detector. A one-way ANOVA with repeated measures model was used for each detector in order to estimate the

between-sessions standard deviation. An estimation of the coefficient of repeatability (COR) can be calculated using equation 2.3. The absolute difference between two measurements on a subject should differ by no more than this value on 95% of occasions [52]. Bland-Altman analysis of the results, following the framework outlined by Bartlett and Frost, was undertaken to visualise any trends or systematic differences between measurement sessions [53].

$$\text{COR} = 1.96 \times \sqrt{2} \times \text{between-sessions SD} \tag{2.3}$$

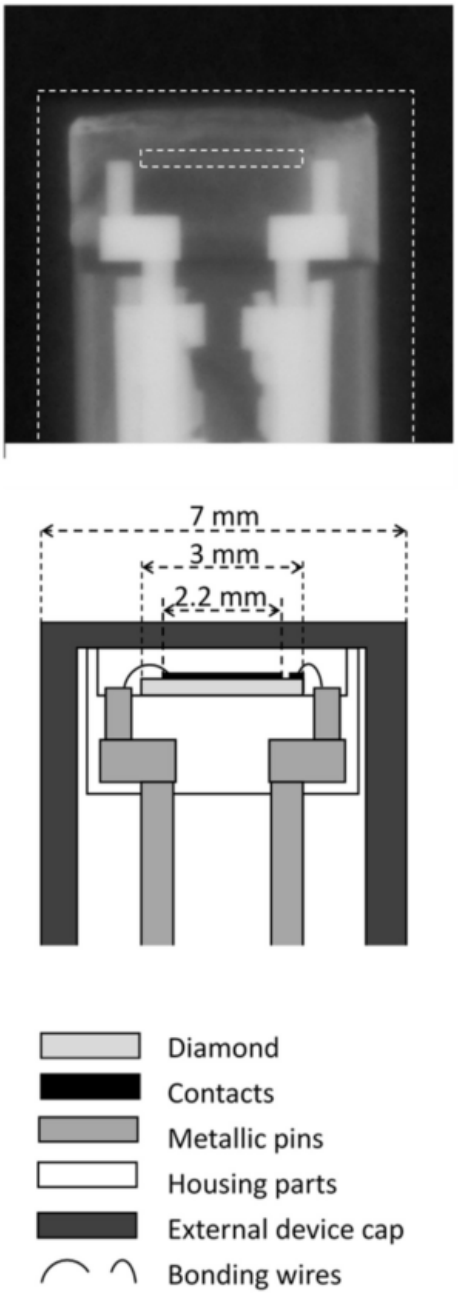


Fig. 2.3 X-ray image and geometrical sketch of microDiamond detector, reproduced from Marinelli et al. 2016 [50].

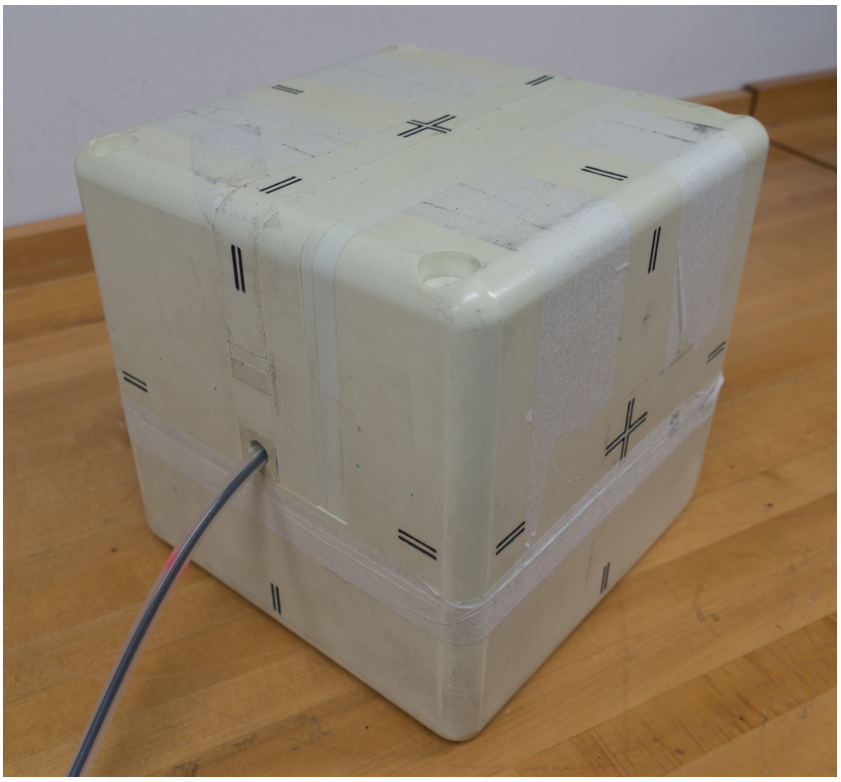


Fig. 2.4 CIRS Model 009 Cube 20 with CC04 chamber positioned at centre of the phantom. Reference marks are used to align centre of phantom with the linac isocentre.

Chapter 3

Results

3.1 Point Dose Results

3.1.1 Repeatability Study

In order to quantify the repeatability of point dose measurements, 10 plans were selected to be remeasured in two additional measurement sessions. Each plan was measured once per measurement session, for a total of three measurements per plan per detector. A summary of repeatability results are shown in table 3.1. The between-sessions standard deviation for each detector was estimated by fitting a repeated-measures ANOVA model. Sphericity was not assumed and a Geisser-Greenhouse correction was used. An assumption of the ANOVA model is that the standard deviation of measurement errors is consistent throughout the range of 'error-free' values. This was verified using Bland-Altman analysis of results between measurement sessions. This qualitatively displays paired differences against population mean, refer to appendices B.1 to B.4. The between-sessions standard deviation was less than 0.5% of the mean measured dose for all detectors excluding the photon diode. Figure 3.1 shows the coefficients of repeatability for each detector as a percentage of mean measured dose. The CC01, CC04, and EFD were the most repeatable, with coefficients of repeatability of 0.29%,

0.4%, 0.55% respectively. The coefficients of repeatability for the microDiamond and PFD were both above 1% (1.30% and 1.53% respectively).

Table 3.1 Repeatability analysis of remeasured plans using all five detectors.

Detector	Between-sessions SD (mGy)	Coefficient of Repeatability (mGy)	95% CI
CC01	2.5	6.9	5.2 - 10
CC04	3.3	9.3	7.0 - 14
EFD	4.6	12.8	9.7 - 19
PFD	12.9	35.9	27 - 53
microDiamond	11.0	30.5	23 - 45

Abbreviations: SD, standard deviation; CI, confidence interval.

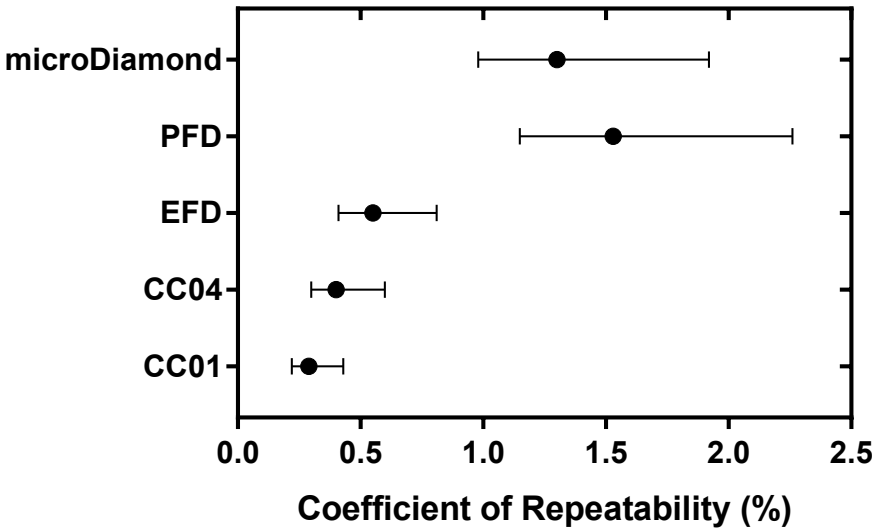


Fig. 3.1 Repeatability Coefficients with corresponding 95% confidence intervals for each detector obtained over 3 measurement sessions, expressed as percentage of mean measured dose.

3.1.2 Treatment Planning System Calculation

Isocentre dose was calculated in Pinnacle (with both 4 degree and 2 degree control point spacing) and RayStation. Post-hoc analysis of a repeated measures ANOVA with Greenhouse-Geisser correction, using Tukey's multiple comparison method was undertaken to investigate pairwise

difference between mean isocentre doses calculated with each beam model. A statistically significant difference in mean isocentre dose for all calculation techniques ($P < 0.0001$) at the 95% confidence level was found.

3.1.3 Dosimeters

Pairwise differences between dosimeter measurements were evaluated using Tukey's multiple comparison method with the results of a repeated measures ANOVA using Greenhouse-Geisser correction. The results are presented in table 3.2. Statistically significant differences in mean dose were found between the CC04 chamber and all other detectors.

Table 3.2 Summary of Tukey's multiple comparison of mean detector measurement results. Bold font indicates statistically significant correlations at 95% confidence level.

	Mean Difference	**P Value**
CC01 vs. CC04	**-0.0141**	**<0.0001**
CC01 vs. EFD	-0.0026	0.9303
CC01 vs. PFD	0.0065	0.2297
CC01 vs. microDiamond	0.0033	0.5898
CC04 vs. EFD	**0.0115**	**0.0214**
CC04 vs. PFD	**0.0206**	**<0.0001**
CC04 vs. microDiamond	**0.0174**	**<0.0001**
EFD vs. PFD	0.0090	0.0689
EFD vs. microDiamond	**0.0058**	**0.0109**
PFD vs. microDiamond	-0.0032	0.8446

Agreement between measured and calculated isocentre point doses can be seen in figure 3.2. A summary of mean point dose agreement can be found in table 3.3. All detectors, aside from the PFD, achieve agreement better than 0.5% for both 2 degree and 4 degree gantry angle spacing in Pinnacle. The EFD displayed the best agreement with Pinnacle; -0.225% and -0.054% for 4 and 2 degree gantry spacing respectively. The disagreement between measured and calculated dose is statistically significant ($P < 0.05$) for all detectors apart from the CC04 and EFD for 4 degree gantry spacing. When 2 degree gantry spacing is used, only the CC01 and EFD have no

statistically significant difference. There exists a systematic offset between RayStation dose calculation results and measurement - disagreement is greater than 0.5% for all detectors.

Table 3.3 Summary of mean point dose discrepancy for each detector and TPS calculation algorithm.

	Mean Difference Measured - Calculated (%)				
	CC01	**CC04**	**EFD**	**PFD**	**microDiamond**
Pinnacle 4 degree GS	-0.3417	0.240	-0.225	-0.594	-0.463
Pinnacle 2 degree GS	-0.170	0.413	-0.054	-0.423	-0.292
RayStation	-1.175	-0.598	-1.059	-1.425	-1.295

3.2 Plan Complexity

Metrics have been calculated in order to quantify plan complexity and investigate the relationship between plan parameters and point dose discrepancy. Complexity metrics previously published in the literature and novel metrics have been calculated for each treatment plan. Pearsons's correlation analysis between complexity metrics and point dose measurement results for the range of detectors has been undertaken to identify the most significantly correlated metrics [54]. Unless otherwise stated, the following correlation analysis has been undertaken for point dose measurement results using the Pinnacle beam model with 2 degree gantry spacing. This is the 'gold standard' calculation method for this study as the plans were originally optimised in Pinnacle and 2 degree gantry spacing has been shown to increase calculation accuracy.

3.2.1 Established Metrics

Results for established complexity metrics for prostate plans are summarised in the descriptive statistics shown in table 3.4. A summary of correlation analysis between established complexity metrics is presented in table 3.5. Only moderate strength correlations exist between plan MU and other metrics, but strong correlations were found for the PMU. The PM metric is very

strongly correlated with MCSv derived metrics, with correlation coefficients of $r = -0.953$ and $r = 0.917$ for MCSv and LTMCSv respectively. Strong correlations exist between CAS, CLS, and MAD metrics and PI. The MFA metric is most strongly correlated with the SAS metrics.

Table 3.4 Summary established complexity metric results for prostate VMAT plans.

	Minimum	**Maximum**	**Mean (± SD)**	**95% CI**
LT	195	564	410 ± 69.6	390 - 431
MCSv	0.287	0.499	0.376 ± 0.0445	0.363 - 0.389
LTMCSv	0.13	0.397	0.233 ± 0.0539	0.217 - 0.249
PI	2.34	7.44	4.25 ± 0.888	3.99 - 4.52
PM	0.236	0.576	0.443 ± 0.0674	0.423 - 0.464
CAS	0.161	0.595	0.411 ± 0.928	0.384 - 0.439
CLS	0.0171	0.169	0.0885 ± 0.0375	0.0773 - 0.0998
MAD	5.11	15.6	11.3 ± 2.38	10.5 - 12
MFA	1880	4810	2930 ± 555	2760 - 3090
SAS10	0.0364	0.156	0.0904 ± 0.0256	0.0827 - 0.0981
SAS20	0.108	0.351	0.238 ± 0.0491	0.223 - 0.253
MU	281	749	532 ± 73.2	510 - 554
PMU	427	619	526 ± 44.5	512 - 539

Abbreviations: SD, standard deviation; CI, confidence interval.

Table 3.5 Pearson correlation coefficients for relationships between established metrics.

	LT	**MCSv**	**LTMCSv**	**PI**	**PM**	**CAS**	**CLS**	**MAD**	**MFA**	**SAS10**	**SAS20**	**MU**	**PMU**
LT		-0.714	-0.787	0.600	0.765	0.601	0.338	0.758	0.061	0.019	0.206	0.244	0.551
MCSv	-0.714		0.925	-0.679	-0.953	-0.840	-0.541	-0.744	0.522	-0.397	-0.619	-0.326	-0.587
LTMCSv	-0.787	0.925		-0.651	-0.917	-0.788	-0.506	-0.762	0.338	-0.334	-0.510	-0.416	-0.623
PI	0.600	-0.679	-0.651		0.804	0.843	0.717	0.719	-0.236	0.541	0.607	0.350	0.663
PM	0.765	-0.953	-0.917	0.804		0.902	0.635	0.859	-0.346	0.416	0.576	0.308	0.656
CAS	0.601	-0.840	-0.788	0.843	0.902		0.825	0.841	-0.485	0.511	0.629	0.255	0.620
CLS	0.338	-0.541	-0.506	0.717	0.635	0.825		0.598	-0.415	0.376	0.371	0.109	0.367
MAD	0.758	-0.744	-0.762	0.719	0.859	0.841	0.598		-0.083	0.149	0.279	0.059	0.484
MFA	0.061	0.522	0.338	-0.236	-0.346	-0.485	-0.415	-0.083		-0.551	-0.686	-0.134	-0.182
SAS10	0.019	-0.397	-0.334	0.541	0.416	0.511	0.376	0.149	-0.551		0.884	0.475	0.469
SAS20	0.206	-0.619	-0.510	0.607	0.576	0.629	0.371	0.279	-0.686	0.884		0.424	0.571
MU	0.244	-0.326	-0.416	0.350	0.308	0.255	0.109	0.059	-0.134	0.475	0.424		0.581
PMU	0.551	-0.587	-0.623	0.663	0.656	0.620	0.367	0.484	-0.182	0.469	0.571	0.581	

Table 3.6 gives a summary of the statistical significance of linear correlations between established complexity metrics and point dose discrepancy results for the range of detectors. As demonstrated in table 3.6 multiple significant correlations exist for both of the ionisation chambers, none for the diodes, and only a single significant correlation with plan MU for the

microDiamond detector. There is a weak correlation for the MCSv metric with CC04 measured results, indicating dose discrepancy increases for decreasing MCSv (an increase in modulation). Three metrics are significantly correlated with both ionisation chambers (CC01 and CC04): PI, CLS, and MAD. An increase in any of these metrics is correlated with increased point dose discrepancy, as displayed in figure 3.3(a-c). In each case the CC04 measurements have greater Pearson correlation coefficients compared to the CC01 measurements: 0.407 vs. 0.373, 0.392 vs. 0.314, and 0.386 vs. 0.350 for PI, CLS, and MAD respectively. A statistically significant correlation exists for MCSv, PM, and CAS for the CC04 point dose measurements only.

Table 3.6 Correlations between point dose discrepancy and established complexity metrics for Pinnacle dose calculation (2 degree gantry spacing). Bold font indicates statistically significant correlations at 95% confidence level.

	CC01		**CC04**		**EFD**		**PFD**		**microDiamond**	
	r	*p value*	*r*	*p value*	*r*	*p value*	*r*	*p value*	*r*	*p value*
LT	0.2176	0.151	0.231	0.1268	-0.1218	0.4254	-0.2008	0.186	-0.2438	0.1066
MCSv	-0.1755	0.2489	**-0.2982**	**0.0466**	0.2471	0.1017	0.03243	0.8325	0.1608	0.2914
LTMCSv	-0.1747	0.2512	-0.2485	0.0997	0.2476	0.101	0.08355	0.5853	0.2256	0.1362
PI	**0.3728**	**0.0117**	**0.407**	**0.0055**	0.08411	0.5828	0.0566	0.7119	0.04659	0.7612
PM	0.2512	0.096	**0.3416**	**0.0217**	-0.1493	0.3277	-0.04536	0.7673	-0.1247	0.4144
CAS	0.2593	0.0855	**0.3653**	**0.0136**	-0.06715	0.6612	0.0415	0.7866	-0.04256	0.7813
CLS	**0.3136**	**0.0359**	**0.3916**	**0.0078**	0.05404	0.7244	0.1624	0.2864	0.1862	0.2206
MAD	**0.3502**	**0.0184**	**0.3863**	**0.0088**	0.002154	0.9888	-0.03485	0.8202	-0.06075	0.6918
MFA	0.0444	0.7721	-0.1564	0.3048	0.2342	0.1215	-0.1891	0.2135	-0.08613	0.5737
SAS2	0.02304	0.8806	0.07516	0.6236	-0.2149	0.1564	0.1313	0.3898	-0.09208	0.5474
SAS5	-0.02585	0.8662	0.02629	0.8639	0.008204	0.9573	0.1578	0.3005	0.1055	0.4904
SAS10	-0.09306	0.5432	0.01025	0.9467	-0.04438	0.7722	0.03654	0.8116	0.01614	0.9162
SAS20	-0.1276	0.4036	0.006724	0.965	-0.1121	0.4635	0.04279	0.7802	-0.08335	0.5862
MU	-0.02901	0.85	-0.01424	0.926	-0.2631	0.0808	-0.09812	0.5214	**-0.296**	**0.0483**
PMU	0.0939	0.5395	0.1132	0.4593	-0.1316	0.3889	-0.06445	0.674	-0.1654	0.2777

3.2.2 Region of interest

No significant correlations exist for any detector for the original LTMCSv metric. But when the metric is evaluated within a region of 1 cm to 4 cm about the isocentre, a statistically significant correlation is observed for the CC04 detector. Evaluation of the metric for the 1 cm ROI produces the strongest correlation coefficient ($r = -0.358$).

Evaluation of LT across the entire treatment field also produces no statistically significant

correlations. Calculation of the metric within regions of 1 cm to 3 cm about the measurement point leads to statistically significant correlations for the EFD detector. With a 1 cm ROI a correlation coefficient of moderate strength ($r = -0.400$) is produced. Correlation strength decreases with increasing ROI dimensions.

Evaluation of the CAS metric within regions of 2 cm to 9 cm about the measurement point results in increased correlation strength compared to the global evaluation for the CC04 detector. The most significant correlation occurs with ROI dimension of 9 cm which gives $r = 0.395$ (global evaluation correlation coefficient $r = 0.365$).

Calculation of the SAS metrics for regions of 1 cm and 2 cm about the measurement point result in strong correlations with ionisation chamber measurements. For example, the SAS20 with a 1 cm ROI gives correlation coefficients of $r = -0.517$ (CC01) and $r = -0.455$ (CC04), but the utility of such strict conditions on the metric is limited; only 3 of 45 plans achieved a non-zero value for this metric. Decreasing the threshold value from 20 mm further reduces the number of non-zero scores.

3.2.3 Additional Metrics

No statistically significant correlations were found for either the MFP or the ratio $\frac{\text{MFP}}{\text{MFA}}$ for any detector. Table 3.7 summarises correlation results for MLC aperture distribution moment analysis. Statistically significant correlations exist for the EFD measured point doses with skewness and kurtosis, giving correlation coefficients of $r = -0.348$ and $r = -0.342$ respectively. Identical analysis of dose rate distribution yielded no statistically significant correlations.

Only the CC04 detector measurements resulted in statistically significant correlations with the blocked fraction metric, as seen in table 3.8. Significant correlations occurred for ROI dimensions of 1 cm to 6 cm. The strongest correlation was found for a ROI of 4 cm, with a correlation coefficient of $r = 0.378$.

Table 3.7 Correlations between point dose discrepancy and MLC aperture distribution moments for Pinnacle dose calculation (2 degree gantry spacing). Bold font indicates statistically significant correlations at 95% confidence level.

	CC01		CC04		EFD		PFD		microDiamond	
	r	*p value*	*r*	*p value*	*r*	*p value*	*r*	*p value*	*r*	*p value*
Mean	0.0444	0.7721	-0.1564	0.3048	0.2342	0.1215	-0.1891	0.2135	-0.08613	0.5737
Variance	0.04172	0.7855	-0.1518	0.3196	0.1738	0.2536	-0.1989	0.1902	-0.1143	0.4547
Skewness	-0.08654	0.5719	0.08205	0.5921	**-0.3487**	**0.0189**	0.1335	0.3818	0.004486	0.9767
Kurtosis	-0.06545	0.6693	0.1028	0.5015	**-0.3419**	**0.0215**	0.123	0.4208	0.01036	0.9462

Table 3.8 Correlations between point dose discrepancy and blocked fraction metric, calculated over 1 cm to 8 cm ROI, for Pinnacle dose calculation (2 degree gantry spacing). Bold font indicates statistically significant correlations at 95% confidence level.

	CC01		CC04		EFD		PFD		microDiamond	
ROI (cm)	*r*	*p value*	*r*	*p value*	*r*	*p value*	*r*	*p value*	*r*	*p value*
1	0.2445	0.1055	**0.3504**	**0.0183**	-0.1053	0.4912	0.005769	0.97	-0.09764	0.5234
2	0.2614	0.0829	**0.3652**	**0.0136**	-0.09723	0.5252	0.01924	0.9002	-0.09899	0.5177
3	0.2728	0.0698	**0.376**	**0.0109**	-0.09018	0.5558	0.01693	0.9121	-0.09507	0.5345
4	0.2672	0.076	**0.3778**	**0.0105**	-0.1057	0.4897	0.01913	0.9008	-0.08075	0.598
5	0.225	0.1374	**0.3515**	**0.0179**	-0.1265	0.4078	0.02201	0.8859	-0.07157	0.6403
6	0.1607	0.2916	**0.3097**	**0.0384**	-0.139	0.3625	0.09769	0.5232	-0.00926	0.9518
7	0.08221	0.5914	0.2654	0.078	-0.196	0.1969	0.1729	0.256	0.04258	0.7812
8	0.01518	0.9212	0.2118	0.1624	-0.2226	0.1416	0.1959	0.1972	0.06801	0.6571

3.2.4 Influence of Calculation Method

The choice of algorithm has a significant influence on calculated dose, as discussed in section 3.1.2. This influences the relationship between point dose discrepancy and complexity metrics. When 4 degree gantry spacing is used for the Pinnacle dose calculation, no established metrics investigated in this study produce statistically significant correlations. Table 3.9 shows correlation significance results when RayStation is used to calculate dose. The only metric with statistically significant correlations is the CLS with Pearson coefficients of $r = 0.336$ and $r = 0.382$, for the CC01 and CC04 ionisation chambers respectively.

Doses were also calculated using an older version of the department's Pinnacle beam model that has been modelled against less reliable small field measurements. Table 3.10 shows correlations found for 2 degree gantry spacing calculation. There exist statistically significant correlations between point dose discrepancy for small field related metrics (MFA, SAS10, and SAS20) that

Table 3.9 Correlations between point dose discrepancy and established complexity metrics for RayStation dose calculation. Bold font indicates statistically significant correlations at 95% confidence level.

	CC01		**CC04**		**EFD**		**PFD**		**microDiamond**	
	r	*p value*	*r*	*p value*	*r*	*p value*	*r*	*p value*	*r*	*p value*
LT	0.148	0.3319	0.151	0.322	-0.1471	0.335	-0.153	0.3158	-0.1898	0.2117
MCSv	-0.1114	0.4662	-0.2133	0.1594	0.2639	0.0799	0.00701	0.9635	0.1544	0.3113
LTMCSv	-0.07789	0.611	-0.1393	0.3616	0.2826	0.06	0.06468	0.6729	0.2253	0.1368
PI	0.2385	0.1146	0.2542	0.092	0.0162	0.9159	0.06207	0.6854	0.04723	0.758
PM	0.2016	0.1841	0.2688	0.0742	-0.161	0.2906	0.008635	0.9551	-0.07892	0.6063
CAS	0.2076	0.1712	0.2881	0.055	-0.0848	0.5797	0.1192	0.4353	0.02231	0.8843
CLS	**0.3357**	**0.0242**	**0.3822**	**0.0096**	0.07217	0.6376	0.2087	0.1689	0.2303	0.1281
MAD	0.1962	0.1964	0.2166	0.153	-0.07326	0.6325	-0.02347	0.8784	-0.05282	0.7304
MFA	0.1262	0.4089	-0.06049	0.693	0.2686	0.0745	-0.08691	0.5703	0.05069	0.7409
SAS2	0.07678	0.6162	0.1173	0.4428	-0.1718	0.2592	0.2106	0.1649	-0.05665	0.7117
SAS5	-0.07992	0.6018	-0.02804	0.8549	-0.0232	0.8798	0.09504	0.5346	0.02083	0.892
SAS10	-0.08943	0.5591	0.007695	0.96	-0.04293	0.7795	0.04415	0.7734	0.01971	0.8978
SAS20	-0.1682	0.2693	-0.038	0.8043	-0.1322	0.3866	0.05777	0.7062	-0.09095	0.5524
MU	0.09762	0.5235	0.1032	0.4998	-0.1786	0.2405	0.03477	0.8206	-0.1819	0.2319
PMU	0.1238	0.4179	0.1328	0.3845	-0.1053	0.4912	-0.00804	0.9582	-0.118	0.4401

do not exist for the new Pinnacle beam model calculation. These correlations are seen across all detectors excluding the PFD for MFA. The strongest correlation is for the CC01 measured dose discrepancy with a correlation coefficient of $r = 0.471$, as seen in figure 3.4. The original beam model is now known to have been based on old small field beam data measured with less suitable small field detectors. This suggests that complexity metric correlation analysis may be useful in identifying deficiencies in beam modelling.

Table 3.10 Correlations between point dose discrepancy and established complexity metrics for original Pinnacle model dose calculation (2 degree gantry spacing). Bold font indicates statistically significant correlations at 95% confidence level.

	CC01		**CC04**		**EFD**		**PFD**		**microDiamond**	
	r	*p value*	*r*	*p value*	*r*	*p value*	*r*	*p value*	*r*	*p value*
LT	0.0158	0.9179	0.0308	0.8406	-0.19	0.2102	-0.212	0.162	-0.252	0.095
MCSv	0.2368	0.1173	0.1396	0.3606	**0.414**	**0.005**	0.2263	0.1349	**0.375**	**0.011**
LTMCSv	0.151	0.3221	0.0923	0.5466	**0.367**	**0.013**	0.2021	0.1831	**0.358**	**0.016**
PI	-0.028	0.8527	0.0038	0.9802	-0.117	0.4459	-0.1	0.5115	-0.124	0.4162
PM	-0.119	0.4363	-0.043	0.7782	**-0.3**	**0.043**	-0.187	0.2195	-0.28	0.0629
CAS	-0.158	0.2989	-0.071	0.6447	-0.261	0.083	-0.117	0.445	-0.219	0.1481
CLS	0.0047	0.9754	0.077	0.6151	-0.096	0.5302	-0.002	0.9872	0.005	0.9741
MAD	0.0466	0.7614	0.0845	0.5811	-0.133	0.3855	-0.104	0.4979	-0.136	0.3731
MFA	**0.472**	**0.001**	**0.321**	**0.032**	**0.434**	**0.003**	0.16	0.2936	**0.3**	**0.045**
SAS2	-0.154	0.3125	-0.116	0.4474	-0.272	0.0708	0.0428	0.78	-0.207	0.1722
SAS5	-0.248	0.1008	-0.216	0.1548	-0.119	0.4361	-0.032	0.8323	-0.107	0.4861
SAS10	**-0.34**	**0.023**	-0.266	0.0774	-0.18	0.2358	-0.131	0.3919	-0.164	0.2805
SAS20	**-0.47**	**0.001**	**-0.38**	**0.011**	-0.292	0.0517	-0.159	0.2971	**-0.31**	**0.04**
MU	-0.014	0.9281	-0.002	0.9916	-0.211	0.1646	-0.033	0.8271	-0.235	0.1201
PMU	-0.135	0.3765	-0.125	0.4152	-0.226	0.1359	-0.167	0.2736	-0.278	0.0643

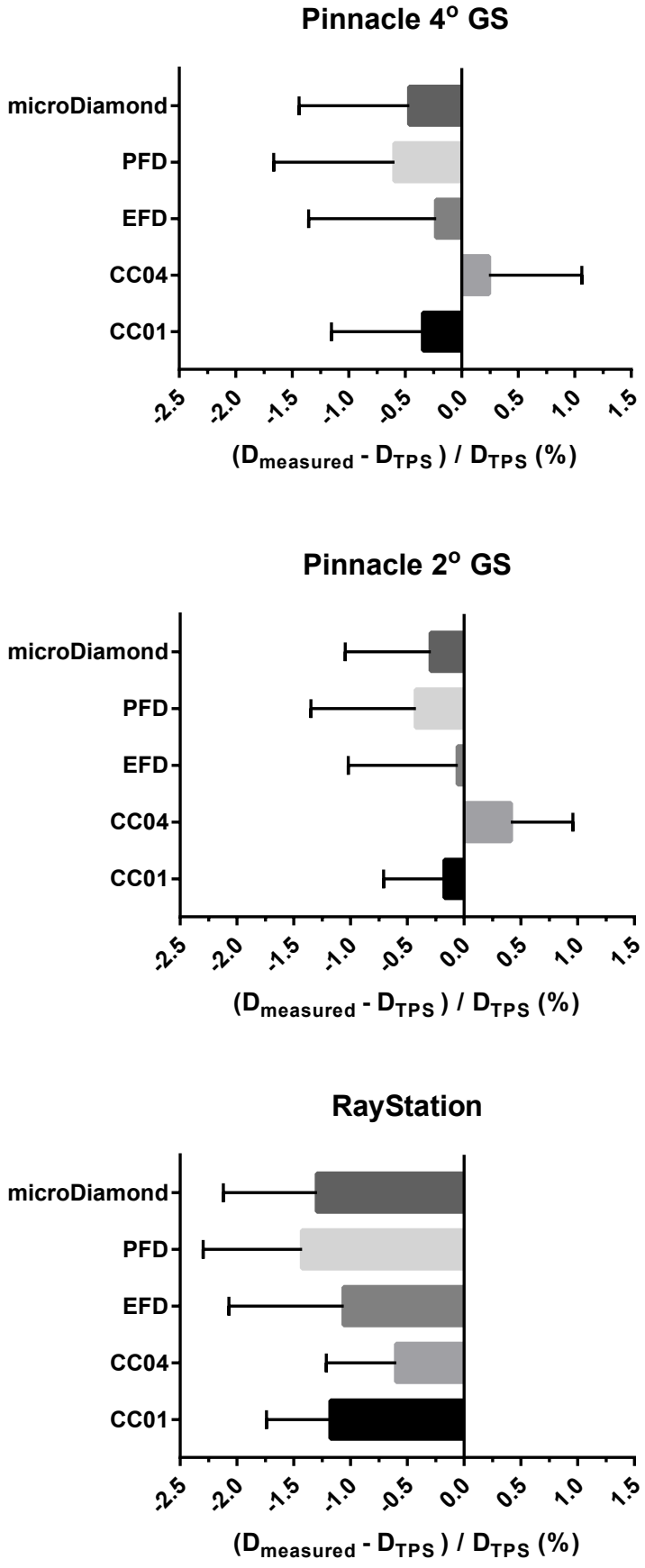


Fig. 3.2 Mean point dose agreement for each calculation method and detector with corresponding standard deviations.

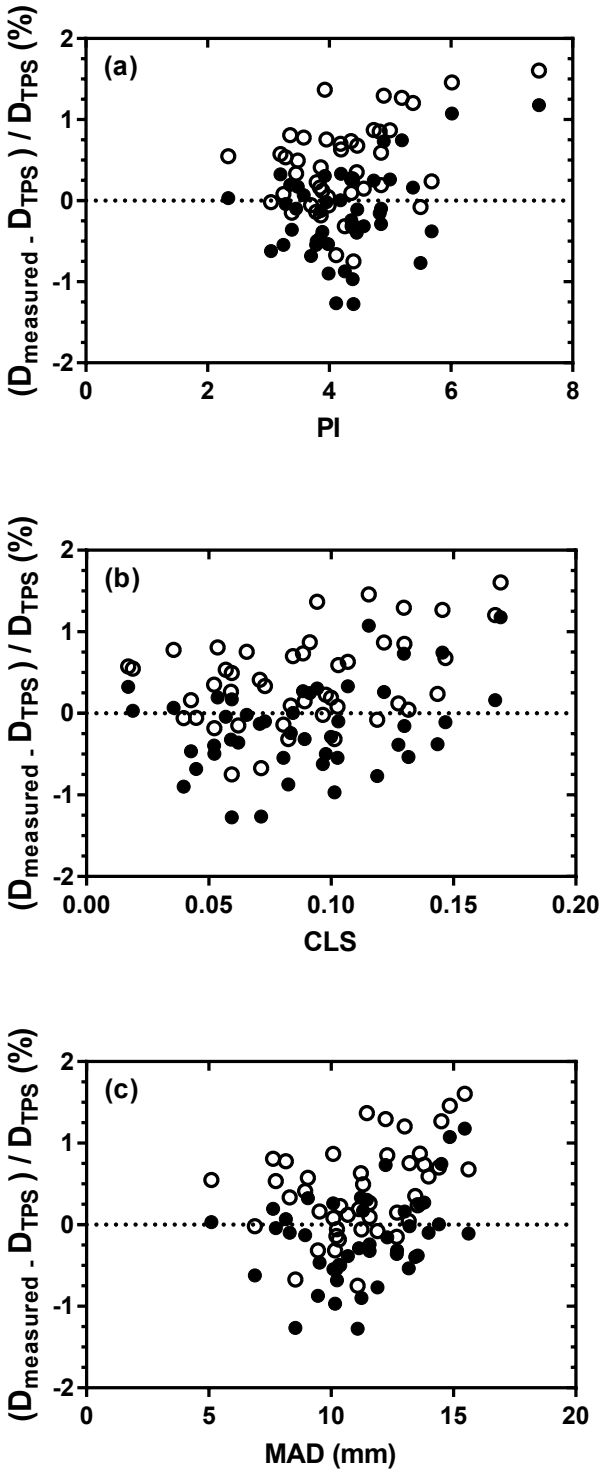


Fig. 3.3 Complexity metrics with statistically significant linear relationships with point dose discrepancy for both CC01 (filled data points) and CC04 (open data points) ionisation chambers. **a** Plan Irregularity (PI), **b** Closed Leaf Score (CLS), **c** Mean Aperture Displacement (MAD). Dose calculated in Pinnacle using 2 degree gantry spacing.

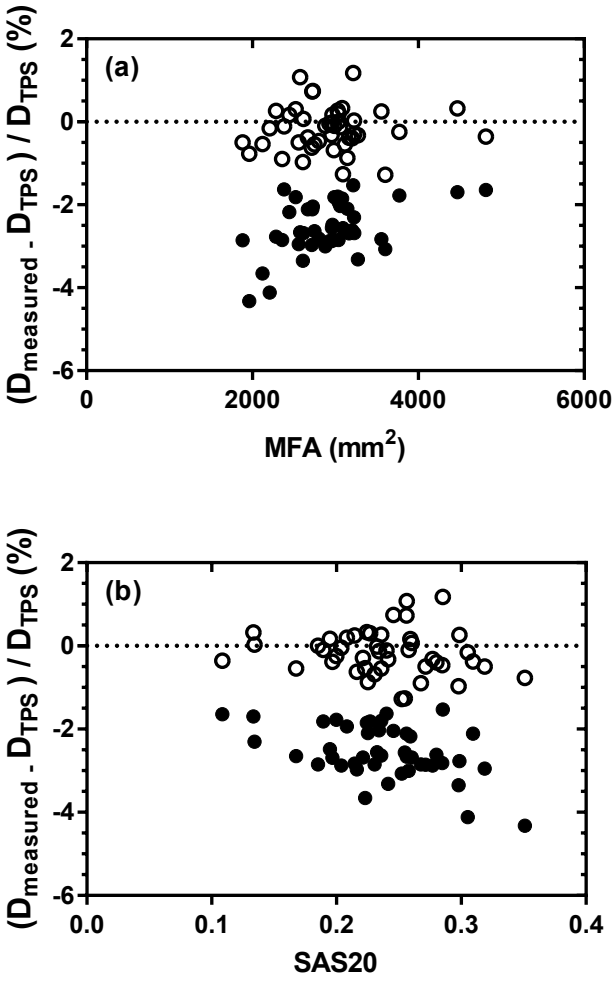


Fig. 3.4 Difference in correlation of small field complexity metrics with point dose discrepancy for both original Pinnacle beam model (filled data points) and new Pinnacle beam model (open data points) for CC01 ionisation chamber. **a** Mean Field Area (MFA) and **b** Small Aperture Score 20 mm (SAS20).

Chapter 4

Discussion

The combination of point dose measurements and gamma analysis of dose plane measurements has been shown to provide the highest sensitivity and specificity in detecting significant dosimetric errors in IMRT [25]. With the shift towards IMRT and VMAT away from 3DCRT, the number of plans requiring verification measurements is unsustainable with current patient specific QA methods. The primary motivation for this project is to identify key plan parameters and understand how they influence point dose discrepancy results in order to optimise the patient specific QA procedure for VMAT treatments.

Ultimately, point dose discrepancy results depend on the accuracy of the measurement and dose calculation. Detector characteristics such as active volume size, energy response, and dose rate response will all influence measured dose. Suitable detector choice is essential to ensure reliable patient specific QA results. In this study five detectors have been investigated.

Treatment planning system dose calculation limitations must be known and evaluated. Calculated dose depends on the choice of calculation algorithm and parameters, such as gantry angle spacing. In this study, treatment plan doses have been calculated in Pinnacle with both 4 degree gantry spacing and interpolated 2 degree gantry spacing. Recalculation of dose in RayStation has also been undertaken. Both measurement and calculation results will be discussed in the first and second sections of this chapter.

The ability to predict which treatment plans are likely to fail patient specific QA *a priori* would reduce physics workload and decrease the impact of any delays due to re-planning of undeliverable treatments. Theoretically this could be achieved with plan complexity metrics which are strongly correlated with plan dosimetric accuracy. Little work has been done to identify complexity metrics which correlate with point dose discrepancy results, with most papers instead focusing on gamma analysis results. The suitability of a range of established complexity metrics and novel ROI based complexity metrics has been investigated. Discussion of these metrics will be undertaken in the third section of this chapter.

4.1 Treatment Planning System Calculation

The following section discusses the influence of planning system calculation technique on dose calculation results. The two treatment planning systems used for VMAT at Auckland City Hospital; Pinnacle and RayStation have been evaluated. In addition, the effect of control point gantry angle spacing for calculated dose is evaluated for the Pinnacle treatment planning system. The correlation of plan complexity metrics and point dose discrepancy has been undertaken in section 3.2.4 and is discussed further in section 4.3.3.

4.1.1 Pinnacle

All of the treatment plans investigated in this study were created in Pinnacle using SmartArc. Plans were optimised and the final dose was calculated using a final gantry spacing of 4 degrees per control point. It has been shown that decreasing the gantry spacing angle leads to increased agreement of calculated field fluence and fluence measured with detector arrays [55]. The choice of 4 degree per control point gantry spacing is used clinically at Auckland City Hospital as a trade off between calculation time and dosimetric accuracy.

Pinnacle allows the interpolation of control points so that plans originally optimised with 4

degree gantry spacing can be calculated with 2 degree gantry spacing. Table 3.3 shows the use of interpolated control points increases the agreement between calculated and measured point doses for all detectors other than the CC04. As discussed in section 2.2.1, Pinnacle approximates VMAT dose delivery as the sum of contributions from static beams at each control point. The agreement between calculation and delivered dose will therefore depend on the degree of leaf modulation between adjacent control points. These results show that moving from the relatively coarse 4 degree gantry spacing to interpolated 2 degree spacing leads to an increase in agreement between calculated and measured dose.

4.1.2 RayStation

A systematic offset between the dose calculated in RayStation measurements result can also be see in table 3.3. All detectors measure point doses lower than those produced by the RayStation calculation. These results are in line with the old Pinnacle beam model investigated in section 3.2.4. As with the old Pinnacle beam model, the RayStation beam model is based on beam data measured using detectors now known to be less suitable for small field dosimetry. Therefore the evaluation of detector performance on the basis of absolute agreement between measurement results and the RayStation calculation has not been undertaken in this study.

4.2 Dosimeters

In this section, the results of point dose measurements are discussed, highlighting the measurement differences between each detector. VMAT treatments produce challenging measurement conditions for detectors, including dynamically changing small fields, and varying dose rates. The energy spectrum and fluence of the beam incident on the detector is variable in a VMAT delivery. Interaction of the beam with MLC leaves positioned about the detector will shift the energy spectrum of the beam reaching the detector when compared to open field conditions.

Charged particle equilibrium cannot be assumed for the whole measurement and perturbations of charged particle fluence by the detector will influence the measured dose. The magnitude of these perturbations is dependent on the geometry of the detector and the material make up of the device. Along with the desirable dosimetric qualities of stability, linearity, dose rate and energy independence, angular independence, and fast response; the ideal detector for VMAT measurements will have a small active volume and be constructed of water equivalent material to reduce the volume averaging effect and fluence perturbation respectively.

It follows that detector choice will influence measurement results significantly, as seen in section 3.1.3. The following sections will discuss the suitability of each detector for use as a point dosimeter in VMAT plan verification measurements.

4.2.1 Repeatability

A random selection of 10 treatment plans were measured three times for each detector over three measurement sessions (15 measurements per plan total, each plan was measured once per session) in order to evaluate the repeatability of point dose measurements for VMAT. Fitting of a repeated measures ANOVA model allowed the estimation of the coefficient of repeatability for each detector. The coefficient of repeatability (also known as the *smallest real difference*) quantifies the absolute measurement error in the same units as the measurement itself. It is the value below which the absolute differences between two measurements would lie with 95% probability [52]. It can be visualised as the 95% limits of agreement used in Bland-Altman analysis that envelope 95% of the differences between repeated measurements on the same subjects.

Three of the five detectors investigated in this study were found to have coefficient of repeatability less than 5 mGy, see table 3.1. The 95% confidence intervals for correlation coefficients overlap for the CC01, CC04 and EFD detectors. Further repeat measurements may allow for sufficient statistical power to differentiate between these detectors, but as it stands all of these

detectors have a coefficient of repeatability less than 0.55% of the mean measured dose. The coefficient of repeatability for the PFD and microDiamond detectors was found to be more than twice this value, at 12.9 mGy and 11.0 mGy respectively. Both ionisation chambers achieved the greatest repeatability, with repeatability coefficient less than 0.5% when expressed as percentage of mean measured value.

The repeatability study design assessed the practical repeatability of point dose measurements which is comprised of delivery repeatability and set-up repeatability together. A recent study investigated VMAT dose delivery reproducibility, and found a relative standard deviation of $< 0.1\%$ between 10 consecutive deliveries [56]. This result suggests the bulk of repeatability error can be attributed to set-up error. Care was taken to select treatment plans where the measurement point was free of steep dose gradients, but there still exists a degree of dose inhomogeneity around the measurement points. To achieve repeatable measurements the position of the detector sensitive volume must be reproducible between measurements sessions.

The smaller the sensitive volume of the detector the more sensitive the detector is to set up errors. The microDiamond detector active volume is effectively a slice 1 μm thick when oriented perpendicularly to the beam. The comparatively larger sensitive volumes of the ionisation chambers provide a volume averaging effect which reduces the impact of small positional offsets. This difference in sensitive volume size may explain why the microDiamond detector coefficient of repeatability is approximately 3 times that of the CC04. The PFD achieved the worst repeatability results, even though the active volume is identical in size to the EFD. The presence of metallic shielding within the PFD encapsulation material may introduce perturbation effects which renders the PFD unsuitable as a detector for VMAT point dose measurements. This effect is discussed further in section 4.2.3.

4.2.2 Ionisation Chambers

Table 3.3 summarises point dose discrepancy results for all detectors and calculation methods. The CC01 chamber measured dose lower than calculated in each case, with the best agreement of -0.170% for the 2 degree gantry spacing Pinnacle calculation. This contrasts with the CC04 measurements which are higher than calculated for both Pinnacle calculations.

Small volume ionisation chambers such as the CC04 are the most commonly used detectors for VMAT and IMRT point dose measurements. Cylindrical ionisation chambers have been well characterised in the literature and have been found to have good stability, linear response to dose, small directional dependence, and insensitivity to beam quality [19]. The most significant challenge for the use of ionisation chambers in VMAT measurements is achieving sufficient spatial resolution. All ionisation chambers produce some degree of volume averaging by virtue of charge collection throughout the active volume of the chamber. Volume averaging has been shown to lead to an under-estimation of dose for output measurements in small fields [57]. To the contrary, in this study the CC04 measurements were higher than calculated. In addition, there exists a statistically significant difference between the CC04 measurements and those for all other detectors, as seen in table 3.2.

A secondary consequence of volume averaging is the blurring of the dose distribution of rapidly spatially varying fields. Modelling the ionisation chamber active volume in the treatment planning system allows the average dose within the volume to be calculated. This method accounts for the volume averaging of the chamber and may lead to better agreement between with measured dose.

The CC01 chamber achieves greater spacial resolution when compared to the CC04 due to a smaller active volume. The central electrode of the CC01 chamber is constructed of steel to increase the ionisation signal, resulting in the sensitivity of the CC01 being approximately 36% of the CC04 with only 25% of the active volume size. Photoelectric interactions within the steel electrode cause the CC01 to over-respond to Compton scatter present in large fields

[58]. Therefore a smaller 4×4 cm^2 reference field for the relative dose calculation may be more optimal than 10×10 cm^2 to avoid underestimation of dose. This over-response in the reference field may account for the lower response of the CC01 chamber when compared to the CC04 which does not over-respond for a 10×10 cm^2 field.

Both the CC01 and CC04 chambers are suitable detectors for point dose measurements in VMAT due to their good repeatability and agreement with treatment planning system dose calculation.

4.2.3 Silicon Diodes

The relative efficiency of per unit volume for silicon diodes is on the order of 10^4 times greater than air ionisation chambers. This allows the construction of diode detectors with very small active volumes while still maintaining high sensitivity. Both diodes investigated in this study have active volumes approximately $1/250^{\text{th}}$ of the CC04 ionisation chamber, but achieve approximately 100 times greater sensitivity. As mentioned earlier, small active volume size reduces volume averaging and is therefore an attractive property for VMAT point dose measurement devices. Traditionally, point dose measurements have been made in regions of homogeneous dose to reduce the impact of volume averaging, but identification of these regions can prove to be difficult where steep dose gradients are required to achieve highly conformal dose distributions. The use of small active volume detectors would allow the measurement of point doses in high dose gradient regions. Kairn et al. have evaluated the suitability of diodes for VMAT QA and found agreement within 2% between diode and ionisation chamber measurements [59].

Table 3.3 summarises the mean point dose discrepancies for each detector and treatment planning system. The EFD achieved the best agreement with dose calculated in Pinnacle for both 4 degree and 2 degree gantry spacing, while the PFD had the worst agreement of any detector with all treatment planning system calculations investigated. As discussed in section

4.2.1, the PFD also has the worst repeatability results of the detectors tested in this study. Conversely, there exists no statistically significant difference for EFD repeatability results and the ionisation chambers.

The construction of both diodes is identical aside from the inclusion of shielding to filter low-energy photons in the PFD. Silicon diode detectors exhibit a higher sensitivity to low-energy photons as a result of the higher atomic number of silicon compared to water. Leading to an over-response in large field sizes, as the proportion of low-energy scattered photons to primary beam photons increases with field size. The addition of shielding within the diode encapsulation is the traditional method to overcome the low-energy over-response. High Z materials are typically used for shielding due to their high photoelectric cross-section. Tungsten powder is used in the encapsulation of the PFD [60]. The presence of metallic shielding within the diode encapsulation is known to be undesirable in small fields and leads to unwanted attenuation of the low-energy spectrum and an over-response from electron scattering [61]. The manufacturer states that the PFD is optimised for large photon fields greater than 10×10 cm^2 [62]. Therefore the PFD may be unsuitable for VMAT point dose measurements, where small field segments are common. The magnitude of variation of detector response for beams entering perpendicularly to the diode axis is approximately 3% for the EFD [63]. Increased directional dependence has been observed for shielded diodes when compared to unshielded diodes [64]. This increased directional dependence may contribute to the poor agreement between measured and calculated point doses using the PFD.

The measurement of a 10×10 cm^2 reference field for dose calculation is unsuitable for the EFD due to the established over-response. A 4×4 cm^2 reference field was selected instead. Kairn et al. recommend that unshielded diodes should not be used for output factor measurements in fields larger than 6×6 cm^2 [65]. The close agreement between EFD measurements and treatment planning system calculations suggest that the over-response of the unshielded diode does not have a significant impact for prostate VMAT plans. The mean jaw defined equivalent

square field size for all plans was 9.1×9.1 cm^2. But when the MLC is taken into account, the average mean field area becomes 29 cm^2, which corresponds to an equivalent square field of $5.4 \times$ cm^2. Although this analysis does not take into account the non-uniform shape of MLC defined fields, therefore the equivalent square field approximation may no longer be strictly valid; the results aid conceptualisation of field size.

4.2.4 Synthetic Diamond

Like silicon diodes, SCD detectors exhibit high radiation sensitivity and small active volume size. Unlike silicon diodes, SCD detectors are nearly tissue equivalent in terms of atomic composition, leading to very similar mass energy absorption coefficients for both water and synthetic diamond. Energy dependence for clinical radiation beams is not significant as the SCD is approximately radiologically equivalent to tissue. SCD detectors also exhibit superior radiation hardness, better temperature stability, and lower leakage currents when compared to silicon diodes.

The above mentioned properties have motivated the dosimetric evaluation of SCD detectors for use in IMRT/VMAT dosimetry. Zani et al. found good agreement between the microDiamond detector and the A1SL ionisation chamber for the measurement of small field output factors, angular dependence of the detector, and the measurement of dose profiles for VMAT treatments of pulmonary disease [66]. Almaviva et al. concluded that for prostate IMRT the microDiamond detector temporal dose response was comparable to small volume ionisation chambers and point dose measurements agree within 2% between the detectors [41]. The point dose measurement results presented in section 3.1.3 are in line with these studies. Similar to all other detectors, aside from the CC04, the microDiamond detector measured point doses lower than calculated for all calculation methods. Measured dose discrepancy was -0.463% and -0.292% respectively for 4 and 2 degree gantry spacing in Pinnacle, demonstrating better agreement than the PFD but worse than the CC01. Like the CC01, a smaller reference field

may have been more suitable for dose calculation; although the magnitude of improvement would be slight due to the near water equivalence of the detector.

Only the PFD detector achieved worse repeatability than the microDiamond detector, shown in table 3.1. The coefficient of repeatability was calculated to be 1.3% of the measured dose. Before each measurement session the microDiamond detector was pre-irradiated with 10 Gy to stabilise dose response to within 0.2%. The within-session dose response repeatability suggests the between-session coefficient of repeatability can be attributed to set-up repeatability. As previously mentioned in section 4.2.1, the smaller the active volume of a detector the greater the sensitivity to set-up errors. The microDiamond detector has the smallest active volume (4×10^{-6} cc) of all detectors investigated in this study. The sensitive volume of the microDiamond detector is only 1 μm thick along the axis of the chamber, resulting in very high spatial resolution. Therefore, small deviations in set up between sessions will effectively result in the measurement of spatially separate point doses, with no overlap due to volume averaging.

4.3 Plan Complexity

This section evaluates the relationship between complexity metrics and point dose discrepancy results, and examines the influence of detector and calculation algorithm choice on the correlation of complexity metrics with point dose discrepancy. Selected plan parameters were characterised via the use of complexity metrics, and point dose measurements were used to evaluate the agreement between calculated and delivered isocentre dose. The complexity metrics were developed under the assumption that increased plan complexity leads to increased uncertainty in VMAT dose calculation and treatment delivery. A selection of established complexity metrics previously published in the literature have been implemented and evaluated against point dose measurement results for all detectors. Novel metrics quantifying MLC perimeter and aperture characteristics, dose rate variation, and the fraction of beam blocked by MLC were also implemented and evaluated. The calculation of plan complexity within a

ROI centred on the measurement point, in order to increase the applicability of complexity metrics for point quality assurance, was investigated for both established and novel metrics. The strength and significance of correlation results for these metrics will be discussed.

Point doses have been calculated for both 4 degree and 2 degree per control point gantry spacing in Pinnacle. Differences in correlations with complexity metrics for each calculation method will be highlighted. The influence of dose calculation algorithm on the relationship between complexity metrics and point dose discrepancy has been investigated.

4.3.1 Established Metrics

The correlation coefficients presented in table 3.5 show significant relationships exist between established complexity metrics. The prostate target volume is roughly spherical about the isocentre; a high proportion of MLC leaves crossing the central axis indicates the presence of irregular fields. Metrics associated with field aperture irregularity (PI, CAS, MAD, and CLS) are strongly correlated as expected. Similarly the strong correlation between the MCSv derived metrics and PM can be explained due to the similarities between calculation of PM and the AAV (one of the two components of the MCSv).The PMU allows the comparison of MU for plans of different prescriptions. The PMU was found to have moderate and strong correlations with all metrics (aside from MFA), confirming the relationship between plan MU and complexity.

The correlation analysis between complexity metrics and point dose discrepancy discussed in this section is based on doses calculated using the Pinnacle algorithm with 2 degree gantry spacing. This represents the 'gold standard' calculation for this study, as discussed in section 3.2. Pearson correlation coefficient calculation results for the established metrics are summarised in table 3.6. There exists no statistically significant correlation for LT, LTMCSv, MFA, SAS, and PMU metrics for any detector investigated. Significant correlations with deliverability metrics, MCSv and PM, are only observed for the CC04 measured results. The direction of the

correlation indicates increased dose discrepancy for increasing modulation.

The strongest correlations for each of the ionisation chambers was seen for the PI metric. Plans consisting of many narrow and/or irregular apertures achieve higher than average PI values. Narrow apertures, like small fields, present challenges for accurate dose measurement due to lateral charged particle disequilibrium [67]. The difficulty of dose calculation in long narrow fields is well documented [11, 68]. A high proportion of long narrow fields leads to an overestimation of output factor [69]. Both ionisation chamber measurement results also show significant correlation with the CLS and MAD accuracy metrics. These metrics quantify the proportion of segments for which deficiencies in MLC modelling, including tongue-and-groove and leaf end effects, could become apparent [70, 71]. A significant relationship between MU and the microDiamond detector was observed. No significant relationships were observed for any of the established complexity metrics and the diodes.

These results suggest that the utility of specific complexity metrics is strongly dependent on detector choice. The cause of the difference in complexity metric correlation results between the ionisation chambers and other detectors is not clear as calculation and measurement inaccuracies are combined in point dose discrepancy analysis.

4.3.2 Localised Complexity and Novel Metrics

The strength of correlations calculated in this study are much lower than has previously been reported in the literature, when metrics are correlated with gamma analysis results [72–74]. Correlation of whole field complexity with dose to a point is problematic as equal weighting is given to all MLC leaves during the complexity calculation. This motivated the investigation of local complexity about the measurement point. The following established metrics were recalculated in square regions of interest about the isocentre, as illustrated in figure 2.1: LT, MCSv, LTMCSv, CAS, CLS, and SAS with thresholds of 2 mm to 20 mm.

Localised complexity analysis may be useful for the evaluation and identification of detector

performance and limitations. When LT is calculated over the whole treatment field, no significant relationships are observed. But when the ROI is reduced to less than 3 cm around the measurement point, significant correlations are observed for the EFD. This is possibly due to the significant energy dependent response of the unshielded diode. Interaction of the beam with MLC leaves causes a shift in energy spectrum of radiation incident on the detector, through MLC leaf scatter and transmission [75]. High leaf travel about the measurement point therefore indicates proportionally higher spectral variation of incident radiation.

The influence of volume averaging for ionisation chamber results was investigated by calculating the BF and SAS metrics with decreasing ROI. Table 3.8 shows significant relationships with BF are only found for the CC04 chamber. Increased correlation strength is observed for SAS calculations with decreased ROI for both the CC01 and CC04 ionisation chambers. This trend is not observed for the other detectors, highlighting the relationship between detector active volume size and volume averaging. The influence of detector occlusion was investigated by the calculation of CAS with 1 cm ROI; no significant increase in the strength of correlations was observed for any detector.

4.3.3 Influence of Calculation Method

As discussed in section 4.1, the choice of calculation algorithm has a significant influence on final calculated dose. The point dose discrepancy calculations, against which complexity metrics have been evaluated in this study, are therefore dependent on calculation algorithm. When 4 degree per control point gantry spacing was used for dose calculation, no statistically significant relationships were found between the established metrics introduced in section 2.3.1 and point dose discrepancy measured with any detector. Masi et al. investigated the impact of control point spacing on gamma pass rates for VMAT by re-optimising high leaf travel plans at 2 and 3 degree gantry spacing [45]. The study showed that correlation strength decreased when using finer control point spacing, leading the authors to conclude the dosimetric accuracy of

high leaf travel plans is improved with reduction of gantry spacing angle. But the Masi et al. study differs from the work presented in this study as the plans were re-optimised with finer gantry spacing rather than recalculated with interpolated control points. The authors justified this approach by calculating mean complexity metric (LT, MCSv, LTMCSv, and MU) values and finding no statistically significant difference between the two plan cohorts.

Recalculation of the Pinnacle optimised treatment plans in RayStation was undertaken to investigate the influence of an entirely independent beam model on the relationship between complexity metrics and dose discrepancy. A summary of the RayStation beam model is presented in section 2.2.2, which is modelled on an independently measured set of beam data. Although there exists a systematic offset between RayStation calculated and measured doses, correlation analysis is still useful to identify significant relationships between plan parameters and point dose discrepancy. Table 3.9 shows that only the CLS metric has a significant relationship with point dose discrepancy. As mentioned in section 4.3.1, the CLS metric identifies calculation conditions which stress the accuracy of MLC modelling. Like the Pinnacle 2 degree gantry spacing results for CLS, significant relationships only exist for the ionisation chamber measurements. Anomalous ionisation chamber response has been observed for measurements made under MLC leaves [76]. The influence of ionisation chamber response in blocked fields over a full VMAT delivery is not significant, due to the very small amount of dose deposited through blocked leaves as a proportion of total dose measured.

Strong correlations with small field complexity metrics (MFA and SAS) for the original Pinnacle beam model can be seen in table 3.10. Statistically significant relationships between point dose discrepancy and MFA exist for both ionisation chambers and the EFD detector. The small field beam data, for which the model was generated, was measured using detectors now known to be less suitable for small field dosimetry. A new set of beam data was measured, using the microDiamond detector following current small field dosimetry guidelines, for the creation of the updated Pinnacle beam model. Table 3.6 shows no significant relationships

were found for small field complexity metrics and dose discrepancy for the new Pinnacle beam model. These results suggest complexity metrics can be used to identify deficiencies in beam modelling and dose calculation algorithm performance.

4.4 Future Work

One limitation of this study was the selection of plans from a single anatomical site. The choice to investigate only prostate treatment plans was motivated by several factors; uniformity of target volume shape between plans, ease of measurement position choice, and the high numbers of clinically delivered prostate plans available for investigation at Auckland City Hospital.
The geometry of the target volume is relatively uniform from plan-to-plan for prostate treatments, when compared to head and neck plans for example. Previous studies have shown complexity metric values are correlated with treatment site, therefore comparison of complexity values between sites is muddied by the intrinsic complexity requirements of different treatment sites. Extension of complexity analysis to other sites would increase the range of complexity metric results, sites associated with both higher and lower complexity should be investigated. For example, head and neck treatment plans have generally been found to require higher modulation to achieve clinical objectives.
Localised complexity analysis within a ROI around the measurement point was introduced to increase the sensitivity of complexity metrics around the measurement point. Modulation of the beam in regions far from the measurement point is included in full-field complexity metric calculations. This modulation is likely to have little influence on the dose delivered to the measurement point, effectively diluting the relationship between complexity metrics and point dose discrepancy. The usefulness of ROI based complexity calculation may increase for treatment plans for sites with larger target volumes, and therefore larger treatment field size.
The selection of measurement point for point dose measurements is crucial. Current best practice is to select a measurement point within a homogeneous high dose region, away from

steep dose gradients [77]. The isocentre is typically a suitable point for prostate VMAT plans. The PTV is roughly circular in shape and relatively uniform between plans, allowing for simple measurement set up without having to apply detector shifts. Positioning of detectors off-axis in solid phantoms can lead to larger set-up inaccuracy [59]. A suitable measurement point, that is in a homogeneous region of high dose, may be unachievable for plans with especially modulated beam fluence where steep dose gradients are present throughout.

Receiver operating characteristic (ROC) analysis was not undertaken to investigate threshold complexity values in order to classify plans as deliverable or un-deliverable. All plans investigated in this study were clinically acceptable and passed quality assurance. The inclusion and analysis of plans which have failed quality assurance may be useful in identification of both delivery and dose calculation algorithm limitations.

A recent paper by Valdes et al. proposed a mathematical framework for the application of machine learning techniques to IMRT/VMAT quality assurance, with the goal of predicting gamma passing rates [78]. The key advantage of adopting a machine learning approach is the integration of many complexity metrics into a single model. The correlation analysis presented in this study has been performed using a single metric at a time, but it is unlikely that a single metric will be able to characterise all of the discrepancy between calculated and measured dose. The determination of each complexity metric weighting in the machine learning algorithm is unique for each beam quality and linear accelerator at each department. This accounts for differences in the relationship between complexity metrics and point dose discrepancy between detectors and calculation algorithms seen in this study. The machine learning based algorithm was able to predict gamma passing rates to within 3% for all 498 IMRT plans analysed. This machine learning approach could be applied to point dose discrepancy results and would allow the identification of key plan parameters and characterise their interaction.

Chapter 5

Conclusion

Patient specific quality assurance is an essential requirement for IMRT and VMAT treatments. Plan verification measurements, including point dose measurements, account for a significant fraction of the overall physics workload. Optimisation of the patient specific QA process is essential as the proportion of patients receiving IMRT and VMAT treatments is increasing.

This study has shown that both the dose calculation algorithm and the choice of detector influence the point dose discrepancy results. A statistically significant difference was found between doses calculated in Pinnacle with 4 degree gantry spacing, which is used clinically, and interpolated 2 degree control points. The suitability of five detectors has been investigated through evaluation of measurement repeatability and agreement with treatment planning system calculations. The ionisation chambers (CC01 and CC04) and unshielded diode (EFD) have been identified as the most suitable detectors for point dose measurements for VMAT patient specific QA. All three detectors achieved excellent repeatability; agreement between measured and calculated dose was highest for the EFD and CC01 detectors.

Plan complexity was quantified through the use of complexity metrics - both established and novel complexity metrics were calculated. Correlation analysis was undertaken to quantify the relationship between point dose discrepancy and plan complexity metrics. The calculation of complexity constrained to a region of interest centred around the measurement and calculation

point increased the sensitivity of some complexity metrics. Detector choice for the point dose discrepancy calculation was shown to influence correlation with complexity metrics. Statistically significant relationships between point dose discrepancy results and complexity metrics were found for only the CC01, CC04, and microDiamond detectors. Likewise, the choice of calculation algorithm is significant; no significant correlations were found when dose was calculated in Pinnacle with 4 degree gantry angle spacing.

Although statistically significant relationships between complexity metrics and point dose discrepancy were found, the strength of these relationships are not sufficient to reliably predict plans likely to fail quality assurance. Therefore, direct measurement of VMAT treatment plans remains an essential component of patient specific quality assurance at Auckland City Hospital.

References

[1] G. Delaney, S. Jacob, C. Featherstone, and M. Barton, "The role of radiotherapy in cancer treatment: estimating optimal utilization from a review of evidence-based clinical guidelines.," *Cancer*, vol. 104, pp. 1129–37, sep 2005.

[2] M. Boadu and M. M. Rehani, "Unintended exposure in radiotherapy: identification of prominent causes.," *Radiotherapy and oncology : journal of the European Society for Therapeutic Radiology and Oncology*, vol. 93, pp. 609–17, dec 2009.

[3] I. El Naqa, P. Pater, and J. Seuntjens, "Monte Carlo role in radiobiological modelling of radiotherapy outcomes.," *Physics in medicine and biology*, vol. 57, no. 11, pp. R75–97, 2012.

[4] F. M. Kahn, *The Physicis of Radiation Therapy*. fourth ed., 2010.

[5] E. B. Podgorsak and K. Kainz, "Radiation Oncology Physics: A Handbook for Teachers and Students," *Medical Physics*, vol. 33, no. 6, p. 1920, 2006.

[6] N. Papanikolaou, T. R. Mackie, C. Meger-Wells, M. Gehring, and P. Reckwerdt, "Investigation of the convolution method for polyenergetic spectra," *Medical Physics*, vol. 20, pp. 1327–1336, sep 1993.

[7] T. R. Mackie, A. F. Bielajew, D. W. Rogers, and J. J. Battista, "Generation of photon energy deposition kernels using the EGS Monte Carlo code.," *Physics in medicine and biology*, vol. 33, pp. 1–20, jan 1988.

[8] T. R. Mackie, J. W. Scrimger, and J. J. Battista, "A convolution method of calculating dose for 15-MV x rays," *Medical Physics*, vol. 12, pp. 188–196, mar 1985.

[9] A. Ahnesjö, "Collapsed cone convolution of radiant energy for photon dose calculation in heterogeneous media," *Medical Physics*, vol. 16, pp. 577–592, jul 1989.

[10] Phillips Medical Systems, "Pinnacle 9.10 Physics Reference Guide," 2014.

[11] RaySearch Laboratories, *RayStation 5: RayPhysics Manual*. 2015.

[12] A. Brahme, J. E. Roos, and I. Lax, "Solution of an integral equation encountered in rotation therapy.," *Physics in medicine and biology*, vol. 27, no. 10, pp. 1221–1229, 1982.

[13] T. Bortfeld, "IMRT: a review and preview.," *Physics in medicine and biology*, vol. 51, no. 13, pp. R363–R379, 2006.

[14] M. Teoh, C. H. Clark, K. Wood, S. Whitaker, and a. Nisbet, "Volumetric modulated arc therapy: a review of current literature and clinical use in practice.," *The British journal of radiology*, vol. 84, no. 1007, pp. 967–96, 2011.

[15] C. X. Yu, "Intensity-modulated arc therapy with dynamic multileaf collimation: an alternative to tomotherapy," *Physics in Medicine and Biology*, vol. 40, pp. 1435–1449, sep 1995.

[16] K. Otto, "Volumetric modulated arc therapy: IMRT in a single gantry arc.," *Medical physics*, vol. 35, pp. 310–7, jan 2008.

[17] D. A. Palma, W. F. Verbakel, K. Otto, and S. Senan, "New developments in arc radiation therapy: A review," *Cancer Treatment Reviews*, vol. 36, pp. 393–399, aug 2010.

[18] J. Yang, G. Tang, P. Zhang, M. Hunt, S. B. Lim, T. LoSasso, and G. Mageras, "Dose calculation for hypofractionated volumetric-modulated arc therapy: approximating continuous arc delivery and tongue-and-groove modeling.," *Journal of applied clinical medical physics*, vol. 17, p. 4989, mar 2016.

[19] D. a. Low, J. M. Moran, J. F. Dempsey, L. Dong, and M. Oldham, "Dosimetry tools and techniques for IMRT.," *Medical physics*, vol. 38, no. 3, pp. 1313–1338, 2011.

[20] D. A. Low, W. B. Harms, S. Mutic, and J. A. Purdy, "A technique for the quantitative evaluation of dose distributions," *Medical Physics*, vol. 25, pp. 656–661, may 1998.

[21] T. Depuydt, A. Van Esch, and D. P. Huyskens, "A quantitative evaluation of IMRT dose distributions: refinement and clinical assessment of the gamma evaluation.," *Radiotherapy and oncology : journal of the European Society for Therapeutic Radiology and Oncology*, vol. 62, pp. 309–19, mar 2002.

[22] B. E. Nelms, M. F. Chan, G. Jarry, M. Lemire, J. Lowden, C. Hampton, M. F. Chan, and J. Lowden, "Evaluating IMRT and VMAT dose accuracy : Practical examples of failure to detect systematic errors when applying a commonly used metric and action levels Evaluating IMRT and VMAT dose accuracy : Practical examples of failure to detect systematic errors wh," vol. 111722, 2013.

[23] B. E. Nelms, H. Zhen, and W. A. Tomé, "Per-beam, planar IMRT QA passing rates do not predict clinically relevant patient dose errors.," *Medical physics*, vol. 38, pp. 1037–44, feb 2011.

[24] M. Stasi, S. Bresciani, A. Miranti, A. Maggio, V. Sapino, and P. Gabriele, "Pretreatment patient-specific IMRT quality assurance: a correlation study between gamma index and patient clinical dose volume histogram.," *Medical physics*, vol. 39, pp. 7626–34, dec 2012.

[25] S. F. Kry, A. Molineu, J. R. Kerns, A. M. Faught, J. Y. Huang, K. B. Pulliam, J. Tonigan, P. Alvarez, F. Stingo, and D. S. Followill, "Institutional Patient-specific IMRT QA Does Not Predict Unacceptable Plan Delivery," *International Journal of Radiation Oncology*Biology*Physics*, vol. 90, pp. 1195–1201, dec 2014.

[26] M. Nauta, J. E. Villarreal-Barajas, and M. Tambasco, "Fractal analysis for assessing the level of modulation of IMRT fields.," *Medical physics*, vol. 38, pp. 5385–93, oct 2011.

[27] A. L. McNiven, M. B. Sharpe, and T. G. Purdie, "A new metric for assessing IMRT modulation complexity and plan deliverability," *Medical Physics*, vol. 37, no. 2, p. 505, 2010.

[28] L. H. Gray, "An Ionization Method for the Absolute Measurement of Formula-Ray Energy," *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 156, pp. 578–596, sep 1936.

[29] L. V. Spencer and F. H. Attix, "A theory of cavity ionization.," *Radiation research*, vol. 3, pp. 239–54, nov 1955.

[30] A. E. Nahum, "Water/air mass stopping power ratios for megavoltage photon and electron beams.," *Physics in medicine and biology*, vol. 23, pp. 24–38, jan 1978.

[31] G. S. Ibbott, "Radiation Dosimetry: Electron Beams with Energies Between 1 and 50 MeV (ICRU Report No. 35)," *Medical Physics*, vol. 12, pp. 813–813, nov 1985.

[32] IAEA, "IAEA TRS-398: Absorbed Dose Determination in External Beam Radiotherapy: An International Code of Practice for Dosimetry based on Standards of Absorbed Dose to Water," tech. rep., 2006.

[33] P. R. Almond, P. J. Biggs, B. M. Coursey, W. F. Hanson, M. S. Huq, R. Nath, and D. W. O. Rogers, "AAPM's TG-51 protocol for clinical reference dosimetry of high-energy photon and electron beams," *Medical Physics*, vol. 26, pp. 1847–1870, sep 1999.

[34] F. Sánchez-Doblado, G. H. Hartmann, J. Pena, R. Capote, M. Paiusco, B. Rhein, A. Leal, and J. I. Lagares, "Uncertainty Estimation in Intensity-Modulated Radiotherapy Absolute Dosimetry Verification," *International Journal of Radiation Oncology Biology Physics*, vol. 68, no. 1, pp. 301–310, 2007.

[35] J. Wuerfel, "Dose measurements in small fields," *Med Phys*, vol. 1, no. 1, pp. 81–90, 2013.

[36] D. a. Low, J. M. Moran, J. F. Dempsey, L. Dong, and M. Oldham, "Dosimetry tools and techniques for IMRT," *Medical Physics*, vol. 38, no. 3, p. 1313, 2011.

[37] J. Seco, B. Clasie, and M. Partridge, "Review on the characteristics of radiation detectors for dosimetry and imaging.," *Physics in medicine and biology*, vol. 59, no. 20, pp. R303–47, 2014.

[38] Ellen Yorke, R. Alecu, L. Ding, D. Fontenla, A. Kalend, D. Kaurin, and M. E. Masterson-McGary, "AAPM Report No. 87: DIODE IN VIVO DOSIMETRY FOR PATIENTS RECEIVING EXTERNAL BEAM RADIATION THERAPY," Tech. Rep. 87, 2005.

[39] G. Rikner and E. Grusell, "General specifications for silicon semiconductors for use in radiation dosimetry.," *Physics in medicine and biology*, vol. 32, no. 9, pp. 1109–1117, 1987.

[40] S. Vatnitsky and H. Järvinen, "Application of a natural diamond detector for the measurement of relative dose distributions in radiotherapy.," *Physics in medicine and biology*, vol. 38, pp. 173–84, jan 1993.

[41] S. Almaviva, I. Ciancaglioni, R. Consorti, F. De Notaristefani, C. Manfredotti, M. Marinelli, E. Milani, A. Petrucci, G. Prestopino, C. Verona, and G. Verona-Rinati, "Synthetic single crystal diamond dosimeters for Intensity Modulated Radiation Therapy applications," *Nuclear Instruments and Methods in Physics Research, Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 608, no. 1, pp. 191–194, 2009.

[42] S. Almaviva, M. Marinelli, E. Milani, G. Prestopino, A. Tucciarone, C. Verona, G. Verona-Rinati, M. Angelone, M. Pillon, I. Dolbnya, K. Sawhney, and N. Tartoni, "Chemical vapor deposition diamond based multilayered radiation detector: Physical analysis of detection properties," *Journal of Applied Physics*, vol. 107, no. 1, 2010.

[43] S. B. Crowe, T. Kairn, N. Middlebrook, B. Sutherland, B. Hill, J. Kenny, C. M. Langton, and J. V. Trapp, "Examination of the properties of IMRT and VMAT beams and evaluation against pre-treatment quality assurance results.," *Physics in medicine and biology*, vol. 60, pp. 2587–2601, mar 2015.

[44] K. Bzdusek, H. Friberger, K. Eriksson, B. Hardemark, D. Robinson, M. Kaus, B. Hardemark, D. Robinson, and M. Kaus, "Development and evaluation of an efficient approach to volumetric arc therapy planning.," *Medical physics*, vol. 36, no. 2009, pp. 2328–2339, 2009.

[45] L. Masi, R. Doro, V. Favuzza, S. Cipressi, and L. Livi, "Impact of plan parameters on the dosimetric accuracy of volumetric modulated arc therapy.," *Medical physics*, vol. 40, no. 7, p. 071718, 2013.

[46] A. L. McNiven, M. B. Sharpe, and T. G. Purdie, "A new metric for assessing IMRT modulation complexity and plan deliverability.," *Medical physics*, vol. 37, no. 2, pp. 505–515, 2010.

[47] W. Du, S. H. Cho, X. Zhang, K. E. Hoffman, and R. J. Kudchadker, "Quantification of beam complexity in intensity-modulated radiation therapy treatment plans.," *Medical physics*, vol. 41, p. 021716, feb 2014.

[48] S. B. Crowe, T. Kairn, J. Kenny, R. T. Knight, B. Hill, C. M. Langton, and J. V. Trapp, "Treatment plan complexity metrics for predicting IMRT pre-treatment quality assurance results.," *Australasian physical & engineering sciences in medicine / supported by the Australasian College of Physical Scientists in Medicine and the Australasian Association of Physical Sciences in Medicine*, vol. 37, pp. 475–82, sep 2014.

[49] I. Ciancaglioni, M. Marinelli, E. Milani, G. Prestopino, C. Verona, G. Verona-Rinati, R. Consorti, a. Petrucci, and F. De Notaristefani, "Dosimetric characterization of a synthetic single crystal diamond detector in clinical radiation therapy small photon beams," *Medical Physics*, vol. 39, no. 7, p. 4493, 2012.

[50] M. Marinelli, G. Prestopino, C. Verona, and G. Verona-Rinati, "Experimental determination of the PTW 60019 microDiamond dosimeter active area and volume," *Medical Physics*, vol. 43, no. 9, pp. 5205–5212, 2016.

[51] R. Ramaseshan, K. Kohli, F. Cao, and R. K. Heaton, "Dosimetric evaluation of Plastic Water Diagnostic-Therapy.," *Journal of applied clinical medical physics / American College of Medical Physics*, vol. 9, no. 2, p. 2761, 2008.

[52] S. Vaz, T. Falkmer, A. E. Passmore, R. Parsons, and P. Andreou, "The Case for Using the Repeatability Coefficient When Calculating Test – Retest Reliability," vol. 8, no. 9, pp. 1–7, 2013.

[53] J. W. Bartlett and C. Frost, "Reliability , repeatability and reproducibility : analysis of measurement errors in continuous variables," no. December 2007, pp. 466–475, 2008.

[54] M. M. Mukaka, "A guide to appropriate use of Correlation coefficient in medical research.," *Malawi Medical Journal*, vol. 24, no. 3, pp. 69–71, 2012.

[55] C. G. Rowbottom, C. Golby, S. Atherton, and R. I. Mackay, "Investigation into the Pinnacle SmartArc Module for VMAT Planning," pp. 721–724, 2009.

[56] T. Satherley, *The impact of plan complexity on the accuracy of VMAT for the treatment of head and neck cancer by*. PhD thesis, University of Canterbury, 2015.

[57] H. Benmakhlouf, J. Sempau, and P. Andreo, "Output correction factors for nine small field detectors in 6 MV radiation therapy photon beams: A PENELOPE Monte Carlo study," *Medical Physics*, vol. 41, p. 041711, apr 2014.

[58] C. Martens, C. De Wagter, and W. De Neve, "The value of the PinPoint ion chamber for characterization of small field segments used in intensity-modulated radiotherapy," *Physics in Medicine and Biology*, vol. 45, no. 9, pp. 2519–2530, 2000.

[59] T. Kairn, S. Ibrahim, E. Inness, S. B. Crowe, and J. V. Trapp, "Suitability of Diodes for Point Dose Measurements in IMRT/VMAT Beams," in *World Congress on Medical Physics and Biomedical Engineering, June 7-12, 2015, Toronto, Canada*, pp. 657–660, 2015.

[60] J. M. Lárraga-Gutiérrez, "Experimental determination of field factors ([Formula: see text]) for small radiotherapy beams using the daisy chain correction method.," *Physics in Medicine & Biology*, vol. 60, no. 15, pp. 5813–31, 2015.

[61] I. Griessbach, M. Lapp, J. Bohsung, G. Gademann, and D. Harder, "Dosimetric characteristics of a new unshielded silicon diode and its application in clinical photon and electron beams," *Medical Physics*, vol. 32, pp. 3750–3754, nov 2005.

[62] Iba Dosimetry GmbH, "Diode detectors: EFD, PFD, SFD, RFD. Technical Description.," 2014.

[63] M. Westermark, J. Arndt, B. Nilsson, and A. Brahme, "Comparative dosimetry in narrow high-energy photon beams," *Physics in Medicine & Biology*, vol. 685, 2000.

[64] G. Rikner and E. Grusell, "Selective shielding of a p-Si detector for quality independence.," *Acta radiologica. Oncology*, vol. 24, no. 1, pp. 65–69, 1984.

[65] T. Kairn, P. H. Charles, G. Cranmer-Sargison, S. B. Crowe, C. M. Langton, D. I. Thwaites, and J. V. Trapp, "Clinical use of diodes and micro-chambers to obtain accurate small field output factor measurements," *Australasian Physical and Engineering Sciences in Medicine*, vol. 38, no. 2, pp. 357–367, 2015.

[66] M. Zani, M. Bucciolini, M. Casati, C. Talamonti, M. Marinelli, G. Prestopino, A. Tonnetti, and G. Verona-Rinati, "A synthetic diamond diode in volumetric modulated arc therapy dosimetry.," *Medical physics*, vol. 40, p. 092103, 2013.

[67] S. Sharma, "Challenges of small photon field dosimetry are still challenging," *Journal of Medical Physics*, vol. 39, no. 3, p. 131, 2014.

[68] A. Ahnesjö and M. M. Aspradakis, "Dose calculations for external photon beams in radiotherapy.," *Physics in medicine and biology*, vol. 44, pp. R99–155, nov 1999.

[69] J. L. Bedford, M. D. R. Thomas, and G. Smyth, "Beam modeling and VMAT performance with the Agility 160-leaf multileaf collimator Bedford Journal of Applied Clinical Medical Physics," vol. 14, no. 2, pp. 172–185, 2013.

[70] L. A. Young, F. Yang, N. Cao, and J. Meyer, "Rounded leaf end modeling in Pinnacle VMAT treatment planning for fixed jaw linacs," *Journal of Applied Clinical Medical Physics*, vol. 17, no. 6, pp. 149–162, 2016.

[71] P. Fenoglietto, B. Laliberté, N. Aillères, O. Riou, J.-B. Dubois, and D. Azria, "Eight years of IMRT quality assurance with ionization chambers and film dosimetry: experience of the Montpellier Comprehensive Cancer Center.," *Radiation oncology (London, England)*, vol. 6, no. 1, p. 85, 2011.

[72] J. M. Park, S.-Y. Park, H. Kim, J. H. Kim, J. Carlson, and S.-J. Ye, "Modulation indices for volumetric modulated arc therapy.," *Physics in medicine and biology*, vol. 59, pp. 7315–40, dec 2014.

[73] J. M. Park, S.-Y. Park, and H. Kim, "Modulation index for VMAT considering both mechanical and dose calculation uncertainties.," *Physics in Medicine & Biology*, vol. 60, no. 18, pp. 7101–25, 2015.

[74] L. Masi, R. Doro, V. Favuzza, S. Cipressi, and L. Livi, "Impact of plan parameters on the dosimetric accuracy of volumetric modulated arc therapy.," *Medical physics*, vol. 40, p. 071718, jul 2013.

[75] J. O. Kim, J. V. Siebers, P. J. Keall, M. R. Arnfield, and R. Mohan, "A Monte Carlo study of radiation transport through multileaf collimators," *Medical Physics*, vol. 28, pp. 2497–2506, dec 2001.

[76] D. González-Castaño, J. Pena, F. Sánchez-Doblado, G. H. Hartmann, F. Gómez, and A. Leal, "The change of response of ionization chambers in the penumbra and transmission regions: Impact for IMRT verification," *Medical and Biological Engineering and Computing*, vol. 46, no. 4, pp. 373–380, 2008.

[77] A. Kumar, G. Mukherjee, G. Yadav, V. Pandey, and K. Bhattacharya, "Optimized point dose measurement: An effective tool for QA in intensity-modulated radiotherapy," *Journal of Medical Physics*, vol. 32, no. 4, p. 156, 2007.

[78] G. Valdes, R. Scheuermann, C. Y. Hung, A. Olszanski, M. Bellerive, and T. D. Solberg, "A mathematical framework for virtual IMRT QA using machine learning," *Medical Physics*, vol. 43, pp. 4323–4334, jun 2016.

Appendix A

Calculation of Complexity Metrics

This section summarises how a selection of established complexity metrics have been calculated in this study, for a more complete treatment please see references [47, 74].

Let $x_{A,ijk}$ represent the position of k$^{\text{th}}$ leaf of j$^{\text{th}}$ control point of i$^{\text{th}}$ beam in MLC bank A.

Let $x_{max,B,ij}$ represent the maximum positional change of MLC positions, that is

$$x_{\max,B,ij} = \max(x_{B,ij}) - \min(x_{B,ij}) \tag{A.1}$$

The leaf sequence variability is given by

$$\begin{aligned} \text{LSV}_{ij} = &\left(\frac{1}{(K-1)x_{\max,B,ij}} \sum_{k=1}^{K-1} x_{\max,B,ij} - |x_{B,ij} - x_{B,ij(k+1)}| \right) \\ &\times \left(\frac{1}{(K-1)x_{\max,A,ij}} \sum_{k=1}^{K-1} x_{\max,A,ij} - |x_{A,ij} - x_{A,ij(k+1)}| \right) \end{aligned} \tag{A.2}$$

where K is the number of exposed MLC leaves per bank.

The aperture area variability is

$$\text{AAV}_{ij} = \frac{\sum_{k=1}^{K}(x_{A,ijk} - x_{B,ijk})}{\sum_{k=1}^{K} \max(x_{A,ik}) - \max(x_{Bik})} \tag{A.3}$$

The modulation complexity score for beam i

$$\mathrm{MCSv}_i = \sum_{j=1}^{J-1} \left(\frac{\mathrm{AAV}_{ij} + \mathrm{AAV}_{i(j+1)}}{2} \right) \left(\frac{\mathrm{LSV}_{ij} + \mathrm{LSV}_{i(j+1)}}{2} \right) \frac{\mathrm{MU}_{ij}}{\mathrm{MU}_i} \tag{A.4}$$

The MCSv for a plan consisting of I beams is given by

$$\mathrm{MCSv}_p = \frac{1}{\mathrm{MU}_p} \sum_{i=1}^{I} \mathrm{MCS}_i \mathrm{MU}_i \tag{A.5}$$

The aperture area is given by

$$\mathrm{AA}_{ij} = \sum_{k=1}^{K} t_k (x_{A,ijk} - x_{B,ijk}) \tag{A.6}$$

Similarly the aperture irregularity is

$$\mathrm{AI}_{ij} = \frac{\mathrm{AP}_{ij}^2}{4\pi \mathrm{AA}_{ij}} \tag{A.7}$$

Averaging the AA over all control points of beam i gives the beam area

$$\mathrm{BA}_i = \frac{1}{\mathrm{MU}_i} \sum_j \mathrm{MU}_{ij} \mathrm{AA}_{ij} \tag{A.8}$$

and the beam irregularity

$$\mathrm{BI}_i = \frac{1}{\mathrm{MU}_i} \sum_j \mathrm{MU}_{ij} \mathrm{AI}_{ij} \tag{A.9}$$

The beam modulation is given by

$$\mathrm{BM}_i = 1 - \frac{1}{\mathrm{MU}_i} \sum_j \frac{\mathrm{MU}_{ij} \mathrm{AA}_{ij}}{U(\mathrm{AA}_{ij})} \tag{A.10}$$

where $U(AA_{ij})$ is the union area of all segments in beam i.

Averaging the beam metrics over the whole treatment plan gives the plan aperture

$$\mathrm{PA} = \frac{1}{\mathrm{MU}_p} \sum_i \mathrm{BA}_i \mathrm{MU}_i \tag{A.11}$$

where MU_p is the total MU for the plan.

Similarly for the plan irregularity

$$PI = \frac{1}{\mathrm{MU}_p} \sum_i \mathrm{BI}_i \mathrm{MU}_i \tag{A.12}$$

and plan modulation

$$\mathrm{PM} = \frac{1}{\mathrm{MU}_p} \sum_i \mathrm{BM}_i \mathrm{MU}_i \tag{A.13}$$

plan normalised MU is defined as

$$\mathrm{PMU} = \frac{\mathrm{MU}_p \times 2\mathrm{Gy}}{d} \tag{A.14}$$

where d is the prescribed dose per fraction

Appendix B

Point Dose Repeatability Measurements

Figures B.1 to B.4 show for each detector across all measurement sessions paired differences are not dependent on their mean, therefore we can conclude the standard deviation of errors is independent of the true 'error-free' value.

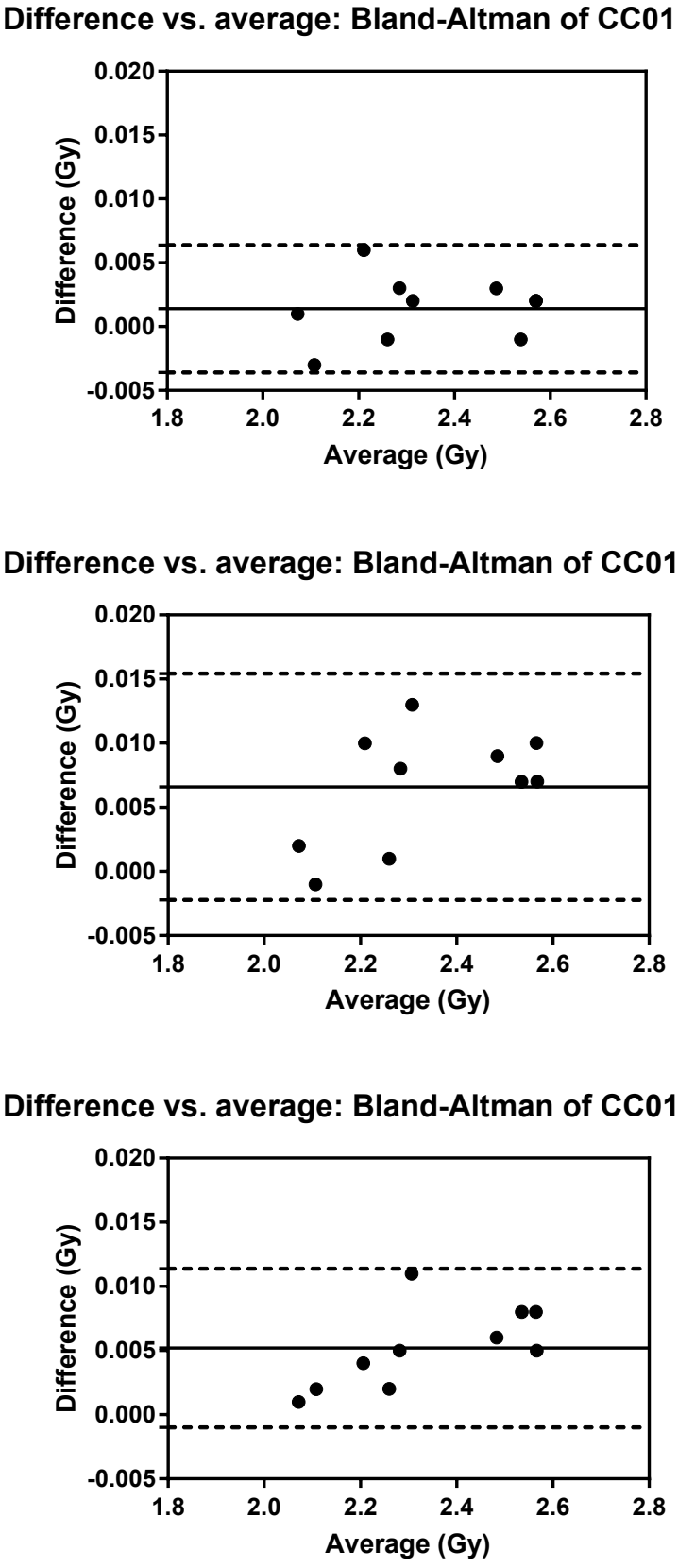


Fig. B.1 Bland-Altman plot of paired differences against mean between measurement sessions for the CC01 detector.

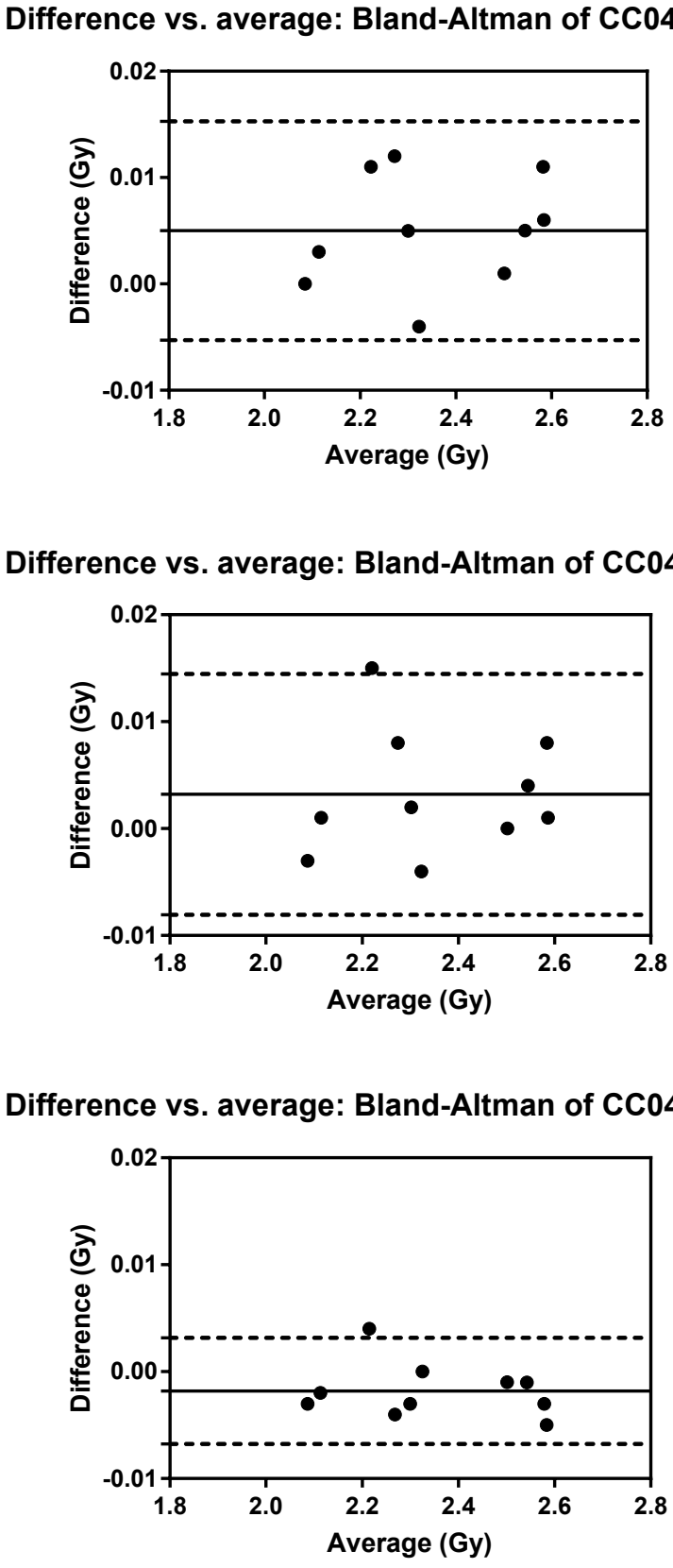
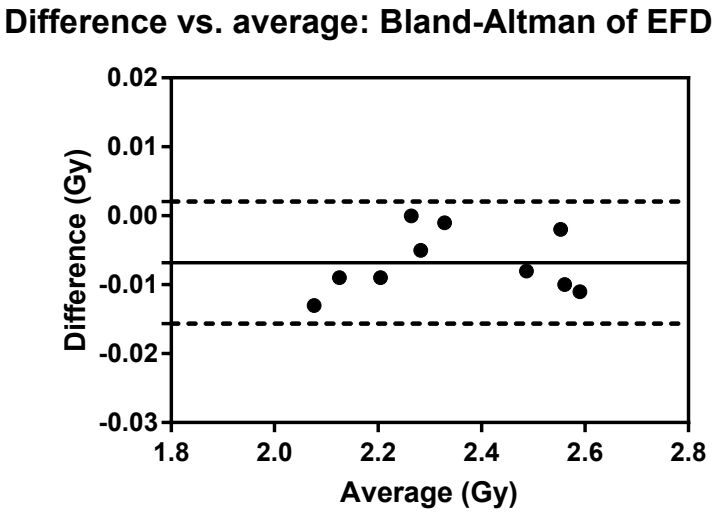
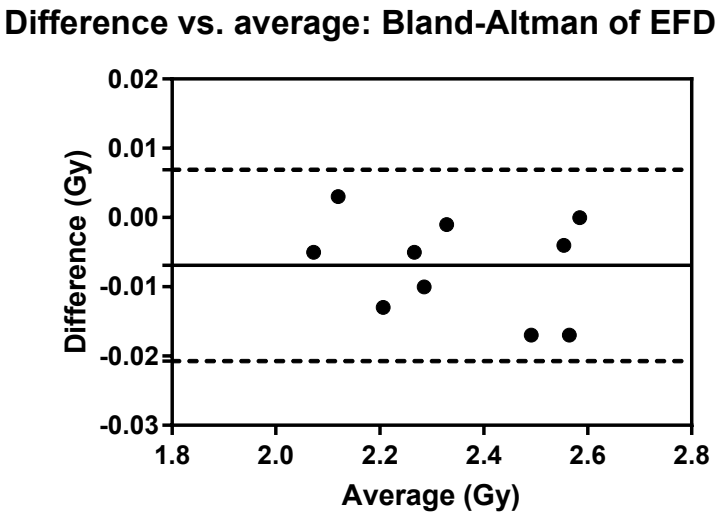


Fig. B.2 Bland-Altman plot of paired differences against mean between measurement sessions for the CC04 detector.





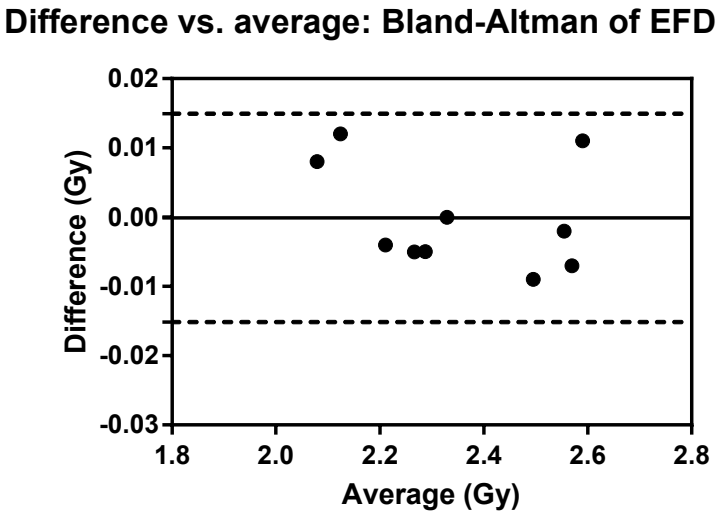
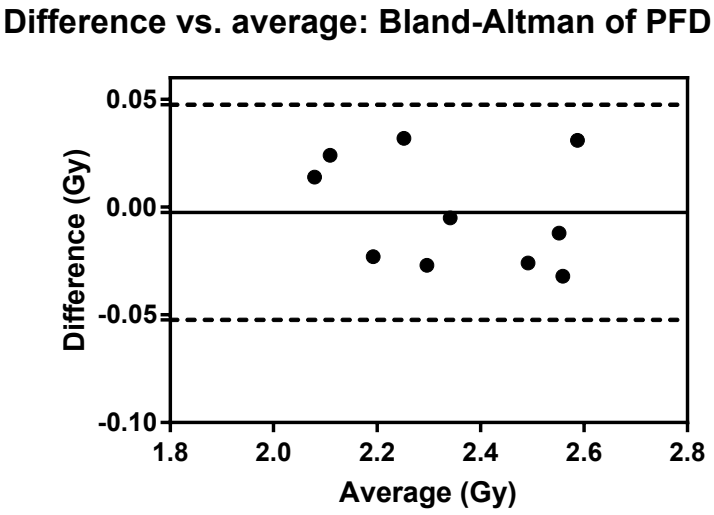
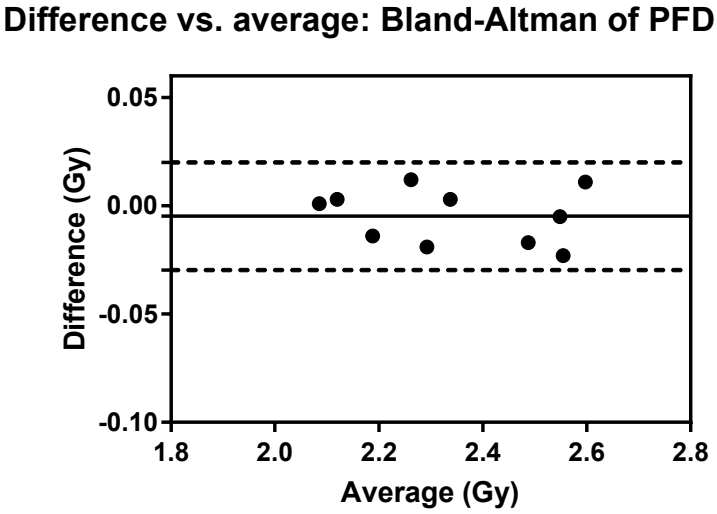


Fig. B.3 Bland-Altman plot of paired differences against mean between measurement sessions for the EFD detector.





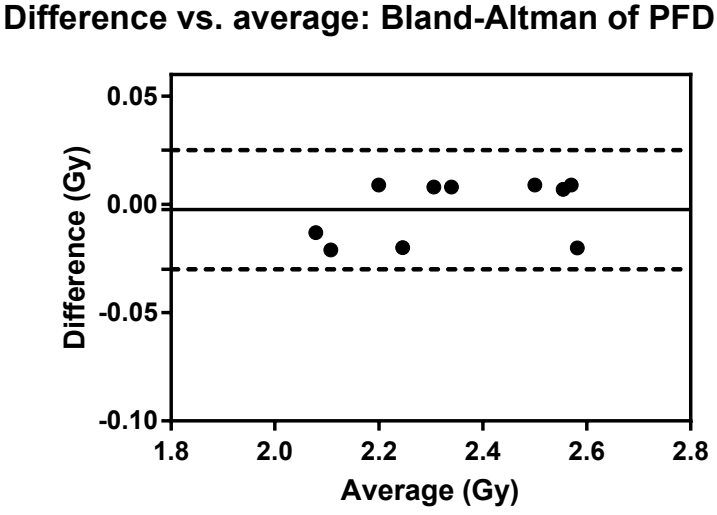


Fig. B.4 Bland-Altman plot of paired differences against mean between measurement sessions for the PFD detector.

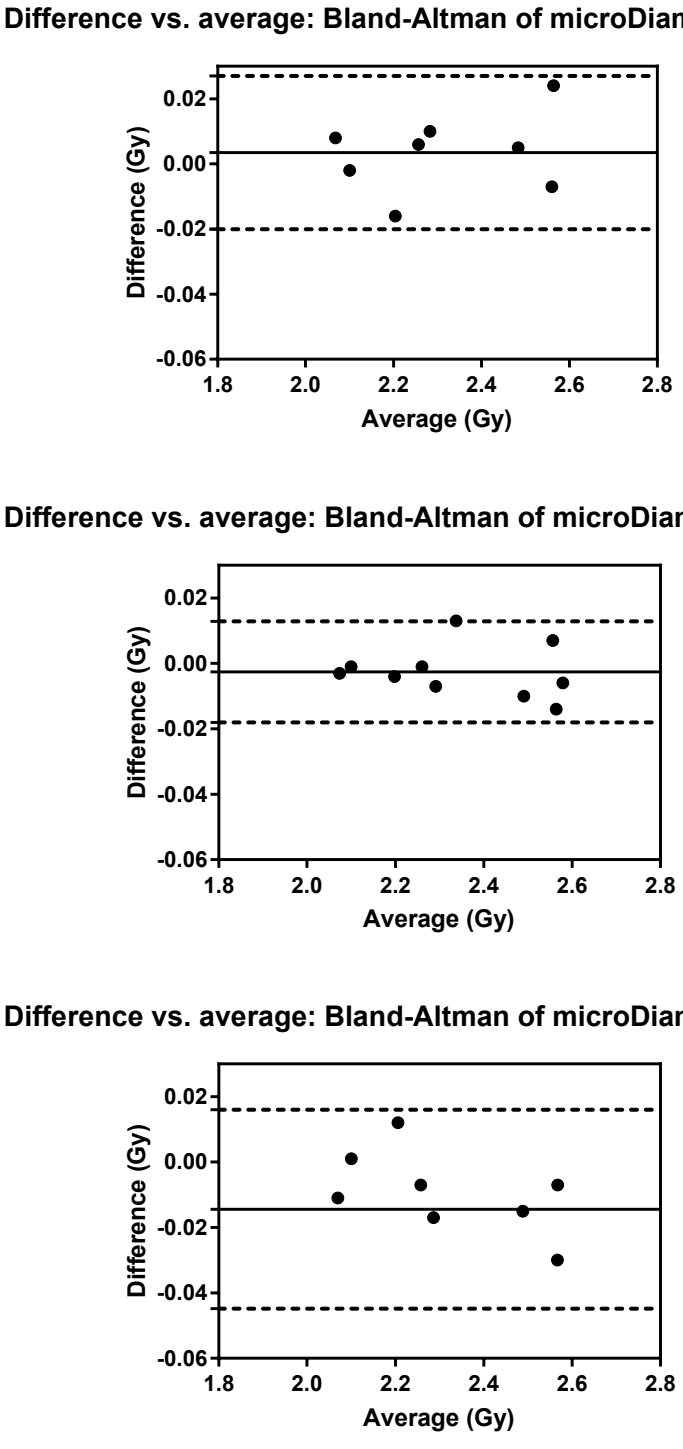


Fig. B.5 Bland-Altman plot of paired differences against mean between measurement sessions for the microDiamond detector.